EQUIDISTRIBUTION OF SPARSE SEQUENCES ON NILMANIFOLDS

NIKOS FRANTZIKINAKIS

Abstract. We study equidistribution properties of nil-orbits $(b^n x)_{n\in\mathbb{N}}$ when the parameter n is restricted to the range of some sparse sequence that is not necessarily polynomial. For example, we show that if $X = G/\Gamma$ is a nilmanifold, $b \in G$ is an ergodic nilrotation, and $c \in \mathbb{R}\setminus\mathbb{Z}$ is positive, then the sequence $(b^{[n^c]}x)_{n\in\mathbb{N}}$ is equidistributed in X for every $x \in X$. This is also the case when n^c is replaced with $a(n)$, where $a(t)$ is a function that belongs to some Hardy field, has polynomial growth, and stays logarithmically away from polynomials, and when it is replaced with a random sequence of integers with sub-exponential growth. Similar results have been established by Boshernitzan when X is the circle.

Contents

1. Introduction and main results

1.1. **Motivation.** A nilmanifold is a homogeneous space $X = G/\Gamma$ where G is a nilpotent Lie group, and Γ is a discrete cocompact subgroup of G. For $b \in G$ and $x = g\Gamma \in X$ we define $bx = (bg)\Gamma$. In recent years it has become clear that studying equidistribution properties of nil-orbits $(b^n x)_{n\in\mathbb{N}}$, and their subsequences, is a central problem, with applications to various areas of mathematics that include combinatorics ([2], [42], [18], [14], [4], [19], [20]), ergodic theory ([11], [28], [41], [29], [34], [16], [17], [30]), number theory ([1], [25], [3], [24]), and probability theory ([13]).

It is well known that for every $b \in G$ and $x \in X$ the sequence $(b^n x)_{n\in\mathbb{N}}$ is equidistributed in some nice algebraic set ([37], [35], [33]), and this is also the case when the parameter n is restricted to the range of some polynomial with integer coefficients ([39], [33]), or the set of prime numbers ([24]). Furthermore, very recently, quantitative equidistribution results for polynomial nil-orbits have been established ([23]) and used as part of an ongoing project to find asymptotics for the number of arithmetic progressions in the set of prime numbers ([22]). These quantitative estimates will also play a crucial role in the present article.

2000 *Mathematics Subject Classification.* Primary: 22F30; Secondary: 37A17.
Key words and phrases. Homogeneous space, nil**man**i**fol**d, equid**is**tribution, Hardy field.
The author was partially supported by NSF grant DMS-0701027.

The main objective of this article is to study equidistribution properties of nil-orbits $(b^nx)_{n\in\mathbb{N}}$ when the parameter n is restricted to some sparse sequence of integers that is not necessarily polynomial. For example, we shall show that if $b \in G$ has a dense orbit in X, meaning $\overline{(b^n\Gamma)}_{n\in\mathbb{N}} = X$, then for every $x \in X$ the sequences

(1)
$$(b^{[n^{\sqrt{3}}]}x)_{n\in\mathbb{N}}, \quad (b^{[n\log n]}x)_{n\in\mathbb{N}}, \quad (b^{[n^2\sqrt{2}+n\sqrt{3}]}x)_{n\in\mathbb{N}}, \quad (b^{[n^3+(\log n)^2]}x)_{n\in\mathbb{N}}, \quad (b^{[(\log(n!))^k]}x)_{n\in\mathbb{N}},$$

are all equidistributed in X. Furthermore, using a probabilistic construction we shall exhibit examples of sequences with super-polynomial growth for which analogous equidistribution results hold (explicit such examples are not known). Let us remark at this point, that since we shall work with sparse sequences of times taken along sequences whose range has typically negligible intersection with the range of polynomial sequences, our results cannot be immediately deduced from known equidistribution results along polynomial sequences.

We shall also study equidistribution properties involving several nil-orbits. For example, suppose that $c_1, c_2, \ldots, c_k$ are distinct non-integer real numbers, all greater than 1, and $b \in G$ has a dense orbit in X. We shall show that the sequence

$$(b^{[n^{c_1}]}x_1, b^{[n^{c_2}]}x_2, \ldots, b^{[n^{c_k}]}x_k)_{n\in\mathbb{N}}$$

is equidistributed in X^k for every $x_1, x_2, \ldots, x_k \in X$.

In a nutshell, our approach is to use the Taylor expansion of a function $a(t)$ to partition the range of the sequence $([a(n)])_{n\in\mathbb{N}}$ into approximate polynomial blocks of fixed degree, in such a way that one can give useful quantitative estimates for the corresponding "Weyl type" sums. In order to carry out this plan, we found it very helpful to deal with the sequence $(a(n))_{n\in\mathbb{N}}$ first, thus leading us to study equidistribution properties of the sequence $(b^{a(n)}x)_{n\in\mathbb{N}}$, where b^s for $s \in \mathbb{R}$ can be defined appropriately.

Before giving the exact results, let us also mention that an additional motivation for our study is the various potential applications in ergodic theory and combinatorics. This direction of research has already proven fruitful; very recently in [20] equidistribution results on nilmanifolds played a key role in establishing a Hardy field refinement of Szemerédi's theorem on arithmetic progressions and related multiple recurrence results in ergodic theory. However, in that article, equidistribution properties involving only conveniently chosen subsequences of the sequences in question were studied. The problem of studying equidistribution properties of the full range of non-polynomial sequences, like those in (1), is more delicate, and is addressed for the first time in the present article. This turns out to be a crucial step towards an in depth study of the limiting behavior of multiple ergodic averages of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{[a_1(n)]}f_1 \cdot \ldots \cdot T^{[a_\ell(n)]}f_\ell,$$

where $(a_i(n))_{n\in\mathbb{N}}$ are real valued sequences that satisfy some regularity conditions. The remaining steps of this project will be completed in a forthcoming paper ([15]).

1.2. **Equidistribution results.** Throughout the article we are going to work with the class of real valued functions $\mathcal{H}$ that belong to some Hardy field (see Section 2.1 for details). Working within the class $\mathcal{H}$ eliminates several technicalities that would otherwise obscure the transparency of our results and the main ideas of their proofs. Furthermore, $\mathcal{H}$ is a rich enough class to enable one to deal, for example, with all the sequences considered in (1).

In various places we evaluate an element b of a connected and simply connected nilpotent Lie group G on some real power s. In Section 2.2 we explain why this operation is legitimate.

When writing $a(t) \prec b(t)$ we mean $a(t)/b(t) \to 0$ as $t \to +\infty$. When writing $a(t) \ll b(t)$ we mean that $|a(t)| \leq C|b(t)|$ for some constant C for every large t. We also say that a function $a(t)$ has *polynomial growth* if $a(t) \prec t^k$ for some $k \in \mathbb{N}$.

1.2.1. *A single nil-orbit.* If $(a(n))_{n\in\mathbb{N}}$ is a sequence of real numbers, and $X = G/\Gamma$ is a nilmanifold, with G connected and simply connected, we say that the sequence $(b^{a(n)}x)_{n\in\mathbb{N}}$ is *equidistributed* in a sub-nilmanifold X_b of X, if for every $F \in C(X)$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F(b^{a(n)}x) = \int F \ dm_{X_b}$$

where m_{X_b} denotes the normalized Haar measure on X_b. Similarly, if the sequence $(a(n))_{n\in\mathbb{N}}$ has integer values, we can define a notion of equidistribution on every nilmanifold X, without imposing any connectedness assumption on G or X.

A sequence $(a(n))_{n\in\mathbb{N}}$ of real numbers is *pointwise good for nilsystems* if for every nilmanifold $X = G/\Gamma$, where G is connected and simply connected, and every $b \in G$, $x \in X$, the sequence $(b^{a(n)}x)_{n\in\mathbb{N}}$ *has a limiting distribution*, meaning, for every $F \in C(X)$ the limit $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F(b^{a(n)}x)$ exists.

We remark that for sequences $(a(n))_{n\in\mathbb{N}}$ with integer values, the connectedness assumptions of the previous definition are superficial. Using the lifting argument of Section 2.2, one sees that if a sequence of integers $(a(n))_{n\in\mathbb{N}}$ is pointwise good for nilsystems, then for every nilmanifold $X = G/\Gamma$, $b \in G$, and $x \in X$, the sequence $(b^{a(n)}x)_{n\in\mathbb{N}}$ has a limiting distribution.

Our first result gives necessary and sufficient conditions for Hardy sequences of polynomial growth to be pointwise good for nilsystems.

Theorem 1.1. *Let $a \in \mathcal{H}$ have polynomial growth.*

Then the sequence $(a(n))_{n\in\mathbb{N}}$ (or the sequence $([a(n)])_{n\in\mathbb{N}}$) is pointwise good for nilsystems if and only if one of the following conditions holds:

- $|a(t) - cp(t)| \succ \log t$ *for every* $c \in \mathbb{R}$ *and every* $p \in \mathbb{Z}[t]$; *or*
- $a(t) - cp(t) \to d$ *for some* $c, d \in \mathbb{R}$ *and some* $p \in \mathbb{Z}[t]$; *or*
- $|a(t) - t/m| \ll \log t$ *for some* $m \in \mathbb{Z}$.

Remarks. • The necessity of these conditions can be seen using rotations on the circle (see [9]). In the case were $X = \mathbb{T}$ their sufficiency was established in [9].

• Unlike the case of integer polynomial sequences, if $p \in \mathbb{R}[t]$, then the sequence $(b^{[p(n)]}x)_{n\in\mathbb{N}}$ may not be equidistributed in a finite union of sub-nilmanifolds of X. For example, when $X = \mathbb{T}(= \mathbb{R}/\mathbb{Z})$, the sequence $(-[n\sqrt{2}]/\sqrt{2}\,\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in the set $\big\{t\mathbb{Z} \colon \{t\} \in [0, 1/\sqrt{2}]\big\}$.

It seems sensible to assert that the first condition in Theorem 1.1 is satisfied by the "typical" function in $\mathcal{H}$ with polynomial growth. It turns out that in this "typical" case, restricting the parameter n of a nil-orbit $(b^n\Gamma)_{n\in\mathbb{N}}$ to the range of the sequence $([a(n)])_{n\in\mathbb{N}}$, does not change its limiting distribution:

Theorem 1.2. *Let $a \in \mathcal{H}$ have polynomial growth and satisfy $|a(t) - cp(t)| \succ \log t$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$.*

(i) If $X = G/\Gamma$ is a nilmanifold, with G connected and simply connected, then for every $b \in G$ and $x \in X$ the sequence $(b^{a(n)}x)_{n\in\mathbb{N}}$ is equidistributed in the nilmanifold $\overline{(b^s x)}_{s\in\mathbb{R}}$.

(ii) If $X = G/\Gamma$ is a nilmanifold, then for every $b \in G$ and $x \in X$ the sequence $(b^{[a(n)]}x)_{n\in\mathbb{N}}$ is equidistributed in the nilmanifold $\overline{(b^n x)}_{n\in\mathbb{N}}$.

Remark. Suppose that we want the conclusion (i) (or (ii)) to be true only for some fixed $b \in G$. Then our proof shows that the assumption can be relaxed to the following: $|a(t) - cp(t)| \succ \log t$ for every $p \in \mathbb{Z}[t]$, and every $c \in \mathbb{R}$ of the form q/β where $q \in \mathbb{Q}$ and β is some non-zero eigenvalue for the nilrotation by b (this means $f(bx) = e(\beta)f(x)$ for some non-constant $f \in L^2(m_X)$). A special case of this stronger result (take $G = \mathbb{R}$, $\Gamma = \mathbb{Z}$, and $b = 1$) gives one of the main results in [8], stating that if $a \in \mathcal{H}$ has polynomial growth and satisfies $|a(t) - p(t)| \succ \log t$ for every $p \in \mathbb{Q}[t]$, then the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$.

1.2.2. *Several nil-orbits.* We give an equidistribution result involving nil-orbits of several Hardy sequences. We say that the functions $a_1(t), \ldots, a_\ell(t)$ have *different growth rates* if the quotient of any two of these functions converges to $\pm\infty$ or to 0.

Theorem 1.3. *Suppose that the functions $a_1(t), \ldots, a_\ell(t)$ belong to the same Hardy field, have different growth rates, and satisfy $t^k \log t \prec a_i(t) \prec t^{k+1}$ for some $k = k_i \in \mathbb{N}$.*

(i) If $X_i = G_i/\Gamma_i$ are nilmanifolds, with G_i connected and simply connected, then for every $b_i \in G_i$ and $x_i \in X_i$, the sequence

$$(b_1^{a_1(n)}x_1, \ldots, b_\ell^{a_\ell(n)}x_\ell)_{n\in\mathbb{N}}$$

is equidistributed in the nilmanifold $\overline{(b_1^s x_1)}_{s\in\mathbb{R}} \times \cdots \times \overline{(b_\ell^s x_\ell)}_{s\in\mathbb{R}}$.

(ii) If $X_i = G_i/\Gamma_i$ are nilmanifolds, then for every $b_i \in G_i$, and $x_i \in X_i$, the sequence

$$(b_1^{[a_1(n)]}x_1, \ldots, b_\ell^{[a_\ell(n)]}x_\ell)_{n\in\mathbb{N}}$$

is equidistributed in the nilmanifold $\overline{(b_1^n x_1)}_{n\in\mathbb{N}} \times \cdots \times \overline{(b_\ell^n x_\ell)}_{n\in\mathbb{N}}$.

Remark. The preceding result contrasts the case of polynomial sequences, where different growth does not imply simultaneous equidistribution for the corresponding nil-orbits. For example, there exists a connected nilmanifold $X = G/\Gamma$ and an ergodic element $b \in G$, such that the sequence $(b^n\Gamma, b^{n^2}\Gamma)_{n\in\mathbb{N}}$ is not even dense in $X \times X$ (see [16]). On the other hand, our result shows that if for instance $a(t) = t^{\sqrt{2}}$, then for every nilmanifold X and ergodic element $b \in G$ the sequence $(b^{[a(n)]}\Gamma, b^{[(a(n))^2]}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in $X \times X$.

It may very well be the case that the hypothesis of Theorem 1.3 can be relaxed to give a much stronger result. The following is a closely related conjecture:

Conjecture. *Let $a_1(t), \ldots, a_\ell(t)$ be functions that belong to the same Hardy field and have polynomial growth. Suppose further that every non-trivial linear combination $a(t)$ of these functions satisfies $|a(t) - cp(t)| \succ \log t$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$.*

Then for every nilmanifold $X = G/\Gamma$, $b_i \in G$, and $x_i \in X$, the sequence

$$(b_1^{[a_1(n)]}x_1, \ldots, b_\ell^{[a_\ell(n)]}x_\ell)_{n\in\mathbb{N}}$$

is equidistributed in the nilmanifold $\overline{(b_1^n x_1)}_{n\in\mathbb{N}} \times \cdots \times \overline{(b_\ell^n x_\ell)}_{n\in\mathbb{N}}$.

1.2.3. *More general classes of functions.* We make some remarks about the extend of the functions our methods cover that do not necessarily belong to some Hardy field.

The conclusions of Theorem 1.2 hold if for some $k \in \mathbb{N}$ the function $a(t)$ is $(k+1)$-times differentiable for large $t \in \mathbb{R}$ and satisfies:

$$a^{(k+1)}(t) \to 0 \text{ monotonically, and } \quad t|a^{(k+1)}(t)| \to \infty. \tag{2}$$

(If $a \in \mathcal{H}$, then (2) is equivalent to "$t^k \log t \prec a(t) \prec t^{k+1}$".) To see this, one can repeat the proof of Theorem 1.3 in this particular setup. More generally, the conclusion of Theorem 1.2

holds for functions $a(t)$ that satisfy the following less restrictive conditions: For some $k \in \mathbb{N}$ the function $a \in C^{k+1}(\mathbb{R}_+)$ satisfies

$$(3) \quad |a^{(k+1)}(t)| \text{ decreases to zero}, \quad 1/t^k \prec a^{(k)}(t) \prec 1, \quad \text{and} \quad (a^{(k+1)}(t))^k \prec (a^{(k)}(t))^{k+1}.$$

(If $a \in \mathcal{H}$, then (3) is equivalent to "$a(t)$ has polynomial growth and $|a(t) - p(t)| \succ \log t$ for every $p \in \mathbb{R}[t]$".) One can see this by repeating verbatim part of the proof of Theorem 1.2. The reader is advised to think of the second condition in (3) as the most important one and the other two as technical necessities (for functions in $\mathcal{H}$ the second condition implies the other two).

Theorem 1.3 can be proved for functions $a_i(t)$ that satisfy condition (2) for some $k_i \in \mathbb{N}$ (call this integer the type of $a_i(t)$), and also, for every $k \in \mathbb{N}$ every non-trivial linear combination of those functions $a_i(t)$ that have type k also satisfies (2).

As for Theorem 1.1, unless one works within a "regular" class of functions like $\mathcal{H}$, it seems hopeless to state a result with explicit necessary and sufficient conditions.

1.2.4. *Random sequences of sub-exponential growth.* So far, we have given examples of sequences that are pointwise good for nilsystems and have polynomial growth. For Hardy sequences of super-polynomial growth, it is indicated in [8] that no growth condition should suffice to guarantee equidistribution on $\mathbb{T}$. On the other hand, explicit sequences of super-polynomial growth like $(e^{\sqrt{n}})_{n\in\mathbb{N}}$ or $(e^{(\log n)^2})_{n\in\mathbb{N}}$ are expected to be pointwise good for nilsystems, but proving this seems to be out of reach at the moment, even for rotations on $\mathbb{T}$.

Nevertheless, using a probabilistic argument, we shall show that there exist very sparsely distributed sequences that are pointwise good for nilsystems. In fact, loosely speaking, we shall see that the only growth condition prohibiting the existence of such examples is exponential growth.

Our probabilistic setup is as follows. Let $(\sigma_n)_{n\in\mathbb{N}}$ be a decreasing sequence of reals in $[0, 1]$. We shall construct random sets of integers by including each integer n in the set with probability $\sigma_n \in [0, 1]$. More formally, let (Ω, Σ, P) be a probability space, and $(X_n)_{n\in\mathbb{N}}$ be a sequence of $0-1$ valued independent random variables with $P(\{\omega \in \Omega\colon X_n(\omega) = 1\}) = \sigma_n$. Given $\omega \in \Omega$ we construct the set of positive integers A^ω by taking $n \in A^\omega$ if and only if $X_n(\omega) = 1$. By writing the elements of A^ω in increasing order we get a sequence $(a_n(\omega))_{n\in\mathbb{N}}$.

If $\sigma_n = 1/n^c$ where $c \in [0, 1)$, then the resulting random sequence has almost surely polynomial growth (in fact it is asymptotic to $n^{1/(1-c)}$). If $\sigma_n = 1/n$, then almost surely the resulting random sequence is bad for pointwise convergence results even for circle rotations (see [31]). Therefore, it makes sense to restrict our attention to the case where $\sigma_n \succ 1/n$. By choosing σ_n appropriately, we can get examples of random sequences with any prescribed sub-exponential growth.

In [7] (and subsequently in [10]) it was shown that if $\lim_{n\to\infty} n\sigma_n = \infty$, then almost surely, the random sequence $(a_n(\omega))_{n\in\mathbb{N}}$ is pointwise good for convergence of rotations on the circle. We extend this result to rotations on nilmanifolds by showing the following:

Theorem 1.4. *Let $(\sigma_n)_{n\in\mathbb{N}}$ be a decreasing sequence of reals satisfying $\lim_{n\to\infty} n\sigma_n = \infty$. Then almost surely, the random sequence $(a_n(\omega))_{n\in\mathbb{N}}$ is pointwise good for nilsystems.*

Remark. As it will become clear from the proof, the condition $\lim_{n\to\infty} n\sigma_n = \infty$ can be replaced with the condition $\lim_{N\to\infty} \frac{\sum_{1\leq n\leq N}\sigma_n}{\log N} = \infty$. Furthermore, our method of proof will show that almost surely, the limits $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N F(b^{a_n(\omega)}x)$ and $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N F(b^n x)$ are equal for every nilmanifold $X = G/\Gamma$, $F \in C(X)$, $b \in G$, and $x \in X$.

1.3. **Applications.** We give some rather straightforward applications of the preceding equidistribution results. We only sketch their proofs leaving some routine details to the reader. For aesthetic reasons, we represent elements $t\mathbb{Z}$ of $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ by t.

The first is an equidistribution result on $\mathbb{T}$, which we do not see how to handle using conventional exponential sum techniques for $k \geq 2$.

Theorem 1.5. *Let $a \in \mathcal{H}$ have polynomial growth and satisfy $|a(t) - cp(t)| \succ \log t$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$.*

Then for every $k \in \mathbb{N}$ and irrational β, the sequence $([a(n)]^k \beta)_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$.

Remark. A standard modification of our argument gives the following more general conclusion: for every $q \in \mathbb{Z}[t]$ non-constant and irrational β, the sequence $\big(q([a(n)])\beta\big)_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$. A similar extension holds for Theorems 1.7, 1.8, and Theorem 1.6 (with ℓ non-constant polynomials).

Proof (Sketch). Suppose for convenience that $k = 2$. We define the transformation $T\colon \mathbb{T}^2 \to \mathbb{T}^2$ by $T(x, y) = \big(x + \beta, y + 2x + \beta\big)$. It is well known that the resulting system is isomorphic to a nilsystem (and the conjugacy map is continuous), and that this system is ergodic if β is irrational. Applying Theorem 1.2 we get that the sequence $(T^{[a(n)]}(0, 0))_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}^2$. An easy computation shows that $T^n(0, 0) = (n\beta, n^2\beta)$, therefore for every $F \in C(\mathbb{T}^2)$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F([a(n)]\beta, [a(n)]^2\beta) = \int F \, dm_{\mathbb{T}^2},$$

where $m_{\mathbb{T}^2}$ denotes the normalized Haar measure on $\mathbb{T}^2$. Using this identity for $F(x, y) = e(ky)$, where k is a non-zero integer, we get

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e(k[a(n)]^2\beta) = 0.$$

This shows that the sequence $([a(n)]^2\beta)_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$. □

Similarly, we can deduce from Theorem 1.3 the following result:

Theorem 1.6. *Let $a_1(t), \ldots, a_\ell(t)$ be functions that belong to the same Hardy field, have different growth rates, and satisfy $t^k \log t \prec a_i(t) \prec t^{k+1}$ for some $k = k_i \in \mathbb{N}$.*

Then for every $l_i \in \mathbb{N}$ and irrationals β_i, the sequence

$$([a_1(n)]^{l_1}\beta_1, \ldots, [a_\ell(n)]^{l_\ell}\beta_\ell)_{n\in\mathbb{N}}$$

is equidistributed in $\mathbb{T}^\ell$.

Next we give an application to ergodic theory. We say that a sequence of integers $(a(n))_{n\in\mathbb{N}}$ is *good for mean convergence*, if for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and $f \in L^2(\mu)$ the averages $\frac{1}{N}\sum_{n=1}^{N} f(T^{a(n)}x)$ converge in $L^2(\mu)$ as $N \to \infty$. Using the spectral theorem for unitary operators, one can see that a sequence $(a(n))_{n\in\mathbb{N}}$ is *good for mean convergence* if and only if for every $t \in \mathbb{R}$ the sequence $(a(n)t)_{n\in\mathbb{N}}$ has a limiting distribution.

Theorem 1.7. *Let $a \in \mathcal{H}$ have polynomial growth and $k \in \mathbb{N}$.*

Then the sequence $([a(n)]^k)_{n\in\mathbb{N}}$ is good for mean convergence if and only if one of the three conditions in Theorem 1.1 is satisfied.

Remark. For $k = 1$ this result was established in [9].

Proof (Sketch). The necessity of the conditions can be seen exactly as in the proof of Theorem 3.1. To prove the sufficiency, we apply Theorem 1.1 for some appropriate unipotent affine transformations of some finite dimensional tori. We deduce that for every $k \in \mathbb{N}$ and $t \in \mathbb{R}$ the sequence $([a(n)]^k t)_{n\in\mathbb{N}}$ has a limiting distribution. As explained before, this implies that the sequence $([a(n)]^k)_{n\in\mathbb{N}}$ is good for mean convergence. □

Lastly, we give a recurrence result for measure preserving systems, and a corresponding combinatorial consequence. We say that a sequence of integers $(a(n))_{n\in\mathbb{N}}$ is *good for recurrence*, if for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and set $A \in \mathcal{B}$ with $\mu(A) > 0$, one has $\mu(A \cap T^{-a(n)}A) > 0$ for some $n \in \mathbb{N}$ such that $a(n) \neq 0$. Using the correspondence principle of Furstenberg ([21]) one can see that this notion is equivalent to the following one: A sequence of integers $(a(n))_{n\in\mathbb{N}}$ is *intersective* if every set of integers Λ with positive upper density contains two distinct elements x and y such that $x - y = a(n)$ for some $n \in \mathbb{N}$.

Theorem 1.8. *Let $a \in \mathcal{H}$ have polynomial growth and satisfy $|a(t) - cp(t)| \succ \log t$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$.*

Then for every $k \in \mathbb{N}$ the sequence $([a(n)]^k)_{n\in\mathbb{N}}$ is good for recurrence (or intersective).

Remarks. • For $k = 1$ this result can be deduced from the equidistribution results in [8].

• A more tedious argument can be used to show that the following weaker assumption suffices: "$a \in \mathcal{H}$ has polynomial growth and satisfies $|a(t) - cp(t)| \to \infty$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$". (By combining the spectral theorem and an argument similar to one used in the proof of Proposition 6.5 in [15], one can handle the case where and $|a(t) - cp(t)| \ll \log t$ for some $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$.)

Proof (Sketch). We apply Theorem 1.2 for some appropriate unipotent affine transformations of finite dimensional tori. We deduce that for every $k \in \mathbb{N}$ and $t \in \mathbb{R}$ the sequence $([a(n)]^k t)_{n\in\mathbb{N}}$ has the same limiting distribution as the sequence $(n^k t)_{n\in\mathbb{N}}$. Using this and the spectral theorem for unitary operators, we conclude that for every invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and set $A \in \mathcal{B}$, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-[a(n)]^k} A) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n^k} A).$$

Since the last limit is known to be positive whenever $\mu(A) > 0$ ([21]), the previous identity shows that the sequence $[a(n)]^k$ is good for recurrence. □

More delicate applications of the equidistribution results presented in Section 1.2 include statements about multiple recurrence and convergence of multiple ergodic averages, and related combinatorial consequences. Such results require much extra work and will be presented in a forthcoming paper ([15]).

1.4. **Structure of the article.** In Section 2 we give the necessary background on Hardy fields and state some equidistribution results on nilmanifolds that will be used later.

In Section 3 we work on a model equidistribution problem that helps us illustrate some of the ideas needed to prove Theorems 1.1 and 1.2. We give a new proof of a result of Boshernitzan on equidistribution of the fractional parts of Hardy sequences of polynomial growth.

In Section 4 we prove Theorems 1.1 and 1.2. The key ingredients are: (i) a reduction step that enables us to "remove" the integer parts and deal with equidistribution properties on nilmanifolds $X = G/\Gamma$ with G connected and simply connected, (ii) the proof technique of the

model problem described in Section 3, and (iii) some quantitative equidistribution results of Green and Tao.

In Section 5 we prove Theorem 1.3. The proof strategy is similar with that of Theorems 1.2, with the exception of a key technical difference that is illustrated using a model equidistribution problem.

In Section 6 we prove Theorem 1.4. We adapt an argument of Bourgain that worked for circle rotations to our more complicated non-Abelian setup.

1.5. **Notational conventions.** The following notation will be used throughout the article: $\mathbb{N} = \{1, 2, \dots\}$, $\mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$, $Tf = f \circ T$, $e(t) = e^{2\pi i t}$, $[t]$ denotes the integer part of t, $\{t\} = t - [t]$, $\|t\| = d(t, \mathbb{Z})$, $\mathbb{E}_{n\in A} a(n) = \frac{1}{|A|}\sum_{n\in A} a(n)$. By $a(t) \prec b(t)$ we mean $\lim_{t\to\infty} a(t)/b(t) = 0$, by $a(t) \sim b(t)$ we mean $\lim_{t\to\infty} a(t)/b(t)$ is a non-zero real number, and by $a(t) \ll b(t)$ we mean $|a(t)| \leq C|b(t)|$ for some constant C for all large t. We use the symbol $\ll_{w_1,\dots,w_k}$ when some expression is majorized by some other expression and the implied constant depends on the parameters $w_1, \dots, w_k$. By $o_{N\to\infty; w_1,\dots,w_k}(1)$ we denote a quantity that goes to zero when the parameters $w_1, \dots, w_k$ are fixed and $N \to \infty$ (when there is no danger of confusion we may omit the parameters). We often write ∞ instead of $+\infty$. For aesthetic reasons, we sometimes represent elements $t\mathbb{Z}$ of $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ by t.

Acknowledgement. We thank the referee for providing constructive comments.

2. Background on Hardy fields and nilmanifolds

2.1. **Hardy fields.** Let B be the collection of equivalence classes of real valued functions defined on some half line (c, ∞), where we identify two functions if they agree eventually.[1] A *Hardy field* is a subfield of the ring $(B, +, \cdot)$ that is closed under differentiation. With $\mathcal{H}$ we denote the *union of all Hardy fields*. If $a \in \mathcal{H}$ is defined in $[1, \infty)$ (one can always choose such a representative of $a(t)$) we call the sequence $(a(n))_{n\in\mathbb{N}}$ a *Hardy sequence*.

A particular example of a Hardy field is the set of all rational functions with real coefficients. Another example is the set $\mathcal{LE}$ that consists of all *logarithmico-exponential functions* ([26], [27]), meaning all functions defined on some half line (c, ∞) by a finite combination of the symbols $+, -, \times, :, \log, \exp$, operating on the real variable t and on real constants. For example the functions $t^{\sqrt{2}}$, $t^2 + t\sqrt{2}$, $t \log t$, e^{t^2}, $e^{\sqrt{\log\log t}}/\log(t^2 + 1)$, are all elements of $\mathcal{LE}$.

We collect here some properties that illustrate the richness of $\mathcal{H}$. More information about Hardy fields can be found in the paper [8] and the references therein.

- $\mathcal{H}$ contains the set $\mathcal{LE}$ and anti-derivatives of elements of $\mathcal{LE}$.
- $\mathcal{H}$ contains several other functions not in $\mathcal{LE}$, like the functions $\Gamma(t)$, $\zeta(t)$, $\sin(1/t)$.
- If $a \in \mathcal{LE}$ and $b \in \mathcal{H}$, then there exists a Hardy field containing both $a(t)$ and $b(t)$.
- If $a \in \mathcal{LE}$, $b \in \mathcal{H}$, and $b(t) \to \infty$, then $a \circ b \in \mathcal{H}$.
  If $a \in \mathcal{LE}$, $b \in \mathcal{H}$, and $a(t) \to \infty$, then $b \circ a \in \mathcal{H}$.
- If a is a continuous function that is algebraic over some Hardy field, then $a \in \mathcal{H}$.

Using these properties it is easy to check that, for example, the sequences $(\log\Gamma(n^2))_{n\in\mathbb{N}}$, $(n^{\sqrt{5}}\zeta(n))_{n\in\mathbb{N}}$, $((\mathrm{Li}(n))^2)_{n\in\mathbb{N}}$ ($\mathrm{Li}(t) = \int_2^t 1/\ln s \, ds$), and the sequences that appear inside the integer parts in (1), are Hardy sequences with polynomial growth. On the other hand, sequences

[1] The equivalence classes just defined are often called *germs of functions*. We choose to use the word function when we refer to elements of B instead, with the understanding that all the operations defined and statements made for elements of B are considered only for sufficiently large values of $t \in \mathbb{R}$.

that oscillate, like $(\sin n)_{n\in\mathbb{N}}$, $(n\sin n)_{n\in\mathbb{N}}$, or the sequence $(e^n+\sin n)_{n\in\mathbb{N}}$ are not Hardy sequences.

We mention some basic properties of elements of $\mathcal{H}$ relevant to our study. Every element of $\mathcal{H}$ has eventually constant sign (since it has a multiplicative inverse). Therefore, if $a\in\mathcal{H}$, then $a(t)$ is eventually monotone (since $a'(t)$ has eventually constant sign), and the limit $\lim_{t\to\infty} a(t)$ exists (possibly infinite). Since for every two functions $a\in\mathcal{H}, b\in\mathcal{LE}$ ($b\neq 0$), we have $a/b\in\mathcal{H}$, it follows that the asymptotic growth ratio $\lim_{t\to\infty} a(t)/b(t)$ exists (possibly infinite). This last property is key, since it will often justify our use of l'Hopital's rule. *We are going to freely use all these properties without any further explanation in the sequel.*

We caution the reader that although every function in $\mathcal{H}$ is asymptotically comparable with every function in $\mathcal{LE}$, some functions in $\mathcal{H}$ are not comparable. This defect of $\mathcal{H}$ will only play a role in one of our results (Theorem 1.3), and can be sidestepped by restricting our attention to functions that belong to the same Hardy field.

A key property of elements of $\mathcal{H}$ with polynomial growth is that we can relate their growth rates with the growth rates of their derivatives:

Lemma 2.1. *Suppose that $a\in\mathcal{H}$ has polynomial growth. We have the following:*

(i) If $t^\varepsilon \prec a(t)$ for some $\varepsilon>0$, then $a'(t)\sim a(t)/t$.

(ii) If $t^{-k}\prec a(t)$ for some $k\in\mathbb{N}$, and $a(t)$ does not converge to a non-zero constant, then $a(t)/(t(\log t)^2)\prec a'(t)\ll a(t)/t$.

Remark. The assumption of polynomial growth is essential, to see this take $a(t)=e^t$. To see that the other assumptions in parts (i) and (ii) are essential take $a(t)=\log t$ for part (i), and $a(t)=e^{-t}$, $a(t)=1+1/t$ for part (ii).

Proof. First we deal with part (i). Applying l'Hopital's rule we get

$$\lim_{t\to\infty}\frac{ta'(t)}{a(t)}=\lim_{t\to\infty}\frac{(\log|a(t)|)'}{(\log t)'}=\lim_{t\to\infty}\frac{\log|a(t)|}{\log t}. \tag{4}$$

Since $t^\varepsilon\prec a(t)$ for some $\varepsilon>0$ and $a(t)$ has polynomial growth, the last limit is a positive real number. Hence, $a'(t)\sim a(t)/t$, proving part (i).

Next we deal with part (ii). First notice that since $a(t)$ does not converge to a non-zero constant we can assume that either $|a(t)|\to\infty$ or $|a(t)|\to 0$.

We show that $a'(t)\ll a(t)/t$. Since $\lim_{t\to\infty}\log|a(t)|=\pm\infty$ we can apply l'Hopital's rule to get (4). Since $a(t)$ has polynomial growth and $t^{-k}\prec a(t)$ for some $k\in\mathbb{N}$, we have that the limit $\lim_{t\to\infty}\log|a(t)|/\log t$ is finite. Using (4) we conclude that the same holds for the limit $\lim_{t\to\infty}(ta'(t))/a(t)$. It follows that $a'(t)\ll a(t)/t$.

Finally we show that $a(t)/(t(\log t)^2)\prec a'(t)$. Equivalently, it suffices to show that the limit

$$\lim_{t\to\infty}\frac{t(\log t)^2a'(t)}{a(t)}$$

is infinite. Arguing by contradiction, suppose this is not the case. Then

$$(\log|a(t)|)'\ll\frac{1}{t(\log t)^2},$$

and integrating we get

$$\log|a(t)|\ll\frac{1}{\log t}+c$$

for some $c\in\mathbb{R}$. It follows that $\log|a(t)|$ is bounded, which contradicts the fact that $|a(t)|\to\infty$ or 0. This completes the proof. $\square$

Following [8], for a non-negative integer k we say that:

(i) The function $a \in \mathcal{H}$ has *type* k if $a(t) \sim t^k$.

(ii) The function $a \in \mathcal{H}$ has *type* k^+ if $t^k \prec a(t) \prec t^{k+1}$.

It is easy to show the following:

Lemma 2.2 (Boshernitzan [8])**.** *Suppose that $a \in \mathcal{H}$ has polynomial growth. Then*

(i) There exists a non-negative integer k such that $a(t)$ has type either k or k^+.

(ii) If $a(t)$ has type k, then $a(t) = ct^k + b(t)$ for some non-zero $c \in \mathbb{R}$ and $b \in \mathcal{H}$ with $b(t) \prec t^k$.

Applying Lemma 2.1 repeatedly we get:

Corollary 2.3. *Suppose that $a \in \mathcal{H}$ has type k^+ for some non-negative integer k.*

Then for every $l \in \mathbb{N}$ with $l \leq k$ we have $a^{(l)}(t) \sim a(t)/t^l$, and for every $l \in \mathbb{N}$ we have

$$a(t)/(t^l(\log t)^2) \prec a^{(l)}(t) \ll a(t)/t^l.$$

Remark. The conclusion fails for some functions of type k. Indeed, if $a(t) = 1 + 1/t$, then $a'(t) \prec a(t)/(t(\log t)^2)$.

2.2. **Nilmanifolds.** Fundamental properties of rotations on nilmanifolds, related to our study, were studied in [1], [36], [35], and [33]. Below we summarize some facts that we shall use, all the proofs can be found or deduced from [33] and [12].

Given a topological group G, we denote its identity element by id_G. By G_0 we denote the connected component of id_G. If $A, B \subset G$, then $[A, B]$ is defined to be the subgroup generated by elements of the form $\{[a, b] : a \in A, b \in B\}$ where $[a, b] = aba^{-1}b^{-1}$. We define the commutator subgroups recursively by $G_1 = G$ and $G_{k+1} = [G, G_k]$. A group G is said to be *nilpotent* if $G_k = \{\mathrm{id}_G\}$ for some $k \in \mathbb{N}$. If G is a nilpotent Lie group and Γ is a discrete cocompact subgroup, then the compact homogeneous space $X = G/\Gamma$ is called a *nilmanifold*. The group G acts on G/Γ by left translation where the translation by a fixed element $b \in G$ is given by $T_b(g\Gamma) = (bg)\Gamma$. We denote by m_X the normalized *Haar measure* on X, meaning, the unique probability measure that is invariant under the action of G by left translations and is defined on the Borel σ-algebra of X. We call the elements of G *nilrotations*. A nilrotation $b \in G$ *acts ergodically on* X, if the sequence $(b^n\Gamma)_{n\in\mathbb{N}}$ is dense in X. When the nilmanifold X is implicit we shall often simply say that a nilrotation $b \in G$ is *ergodic*. It can be shown that if $b \in G$ is ergodic, then for every $x \in X$ the sequence $(b^n x)_{n\in\mathbb{N}}$ is equidistributed in X. A nilrotation $b \in G$ is *totally ergodic*, if for every $r \in \mathbb{N}$ the nilrotation b^r is ergodic. If the nilmanifold X is connected it can be shown that every ergodic nilrotation is in fact totally ergodic.

Example 2.4 (Heisenberg nilmanifold)**.** Let G be the nilpotent group that consists of all upper triangular matrices of the form $\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$ with real entries. If we only allow integer entries we get a subgroup Γ of G that is discrete and cocompact. Then G/Γ is a nilmanifold. It can be shown that a nilrotation $b = \begin{pmatrix} 1 & \alpha & \gamma \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{pmatrix}$ is ergodic if and only if the numbers 1, α, and β are rationally independent.

Let G be a connected and simply connected Lie group and $\exp\colon \mathfrak{g} \to G$ be the exponential map, where $\mathfrak{g}$ is the Lie algebra of G. Since G is a connected and simply connected nilpotent Lie group, it is well known that the exponential map is a bijection. For $b \in G$ and $s \in \mathbb{R}$ we define the element b^s of G as follows: If $X \in \mathfrak{g}$ is such that $\exp(X) = b$, then $b^s = \exp(sX)$.

A more intuitive way to make sense of the element b^s is by thinking of G as a matrix group; then b^s is the element one gets after replacing n by s in the formula giving the elements of the matrix b^n. It is instructive to compare the two equivalent ways of defining b^s in the following example.

Example 2.5 (Heisenberg nilflow)**.** Let X be the Heisenberg nilmanifold. Then the exponential map is given by $\exp\begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & x & z+\frac{1}{2}xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$. As a consequence, if $b = \begin{pmatrix} 1 & \alpha & \gamma \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{pmatrix}$, then $\exp\begin{pmatrix} 0 & \alpha & \gamma-\frac{1}{2}\alpha\beta \\ 0 & 0 & \beta \\ 0 & 0 & 0 \end{pmatrix} = b$, and a short computation shows that $b^s = \begin{pmatrix} 1 & s\alpha & s\gamma+\frac{s(s-1)}{2}\alpha\beta \\ 0 & 1 & s\beta \\ 0 & 0 & 1 \end{pmatrix}$. Alternatively, one can find the same formula for b^s after replacing n by s in the formula $b^n = \begin{pmatrix} 1 & n\alpha & n\gamma+\frac{n(n-1)}{2}\alpha\beta \\ 0 & 1 & n\beta \\ 0 & 0 & 1 \end{pmatrix}$.

Next we record some basic facts that we will frequently use:

(*Basic properties of* b^s). If G is a connected and simply connected Lie group, then for $b \in G$ the map $s \to b^s$ is continuous, and for $s, s_1, s_2 \in \mathbb{R}$ one has the identities $b^{s_1+s_2} = b^{s_1} \cdot b^{s_2}$, $(b^{s_1})^{s_2} = b^{s_1 s_2}$, and $(gbg^{-1})^s = gb^s g^{-1}$.

(*Ratner's theorem, nilpotent case*) Let $X = G/\Gamma$ be a nilmanifold. Then for every $b \in G$ the set $X_b = \overline{\{b^n\Gamma, n \in \mathbb{N}\}}$ has the form H/Δ, where H is a closed subgroup of G that contains b, and $\Delta = H \cap \Gamma$ is a discrete cocompact subgroup of H. Furthermore, the sequence $(b^n\Gamma)_{n\in\mathbb{N}}$ is equidistributed in X_b.

Likewise, if G is connected and simply connected, and $b \in G$, let $Y_b = \overline{\{b^s\Gamma, s \in \mathbb{R}\}}$. Then Y_b has the form H/Δ where H is a closed connected and simply connected subgroup of G that contains all elements b^s for $s \in \mathbb{R}$, and Δ is a discrete cocompact subgroup of H. Furthermore, the nilflow $(b^s\Gamma)_{s\in\mathbb{R}}$ is equidistributed in Y_b.

(*Change of base point formula*). Let $X = G/\Gamma$ be a nilmanifold. As mentioned before, for every $b \in G$ the nil-orbit $(b^n\Gamma)_{n\in\mathbb{N}}$ is equidistributed in the set $X_b = \overline{\{b^n\Gamma, n \in \mathbb{N}\}}$. Using the identity $b^n g = g(g^{-1}bg)^n$ we see that the nil-orbit $(b^n g\Gamma)_{n\in\mathbb{N}}$ is equidistributed in the set $g \cdot X_{g^{-1}bg}$. A similar formula holds when G is connected and simply connected and we replace the integer parameter n with the real parameter s and the nilmanifold X_b with Y_b.

(*Lifting argument*). In several instances it will be convenient for us to assume that a nilmanifold X has a representation G/Γ with G connected and simply connected. To get this extra assumption we argue as follows (see [33]): Since all our results deal with the action on X of finitely many elements of G we conclude that for the purposes of this paper, we can, and will always assume that the discrete group G/G_0 is finitely generated. In this case, one can show that $X = G/\Gamma$ is isomorphic to a sub-nilmanifold of a nilmanifold $\tilde{X} = \tilde{G}/\tilde{\Gamma}$, where $\tilde{G}$ is a connected and simply-connected nilpotent Lie group, with all translations from G "represented" in $\tilde{G}$ (for example if $X = \mathbb{T}$ then $\tilde{X} = \mathbb{R}/\mathbb{Z}$, and if $X = (\mathbb{Z} \times \mathbb{R}^2)/\mathbb{Z}^3$ then $\tilde{X} = \mathbb{R}^3/\mathbb{Z}^3$). Practically, this means that for every $F \in C(X)$, $b \in G$, and $x \in X$, there exists $\tilde{F} \in C(\tilde{X})$, $\tilde{b} \in \tilde{G}$, and $\tilde{x} \in \tilde{X}$, such that $F(b^n x) = \tilde{F}(\tilde{b}^n \tilde{x})$ for every $n \in \mathbb{N}$.

One should keep in mind though when using this lifting trick, that any assumption made about a nilrotation b acting on a nilmanifold X, is typically lost when passing to the lifted nilmanifold $\tilde{X}$. Therefore, the above mentioned construction will be helpful only when our working assumptions impose no restrictions on a nilrotation.

Example 2.6. Let G be the non-connected nilpotent group that consists of all upper triangular matrices of the form $\begin{pmatrix} 1 & k & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$ where $k \in \mathbb{Z}$ and $y, z \in \mathbb{R}$. If we also restrict the entries y and

z to be integers we get a subgroup Γ of G that is discrete and cocompact. In this case, the Heisenberg nilmanifold of Example 2.4 can serve as the lifting $\tilde{X}$ of the nilmanifold $X = G/\Gamma$.

2.3. **Equidistribution on nilmanifolds.** We gather some equidistribution results of polynomial sequences on nilmanifolds that will be used later.

2.3.1. *Qualitative equidistribution on nilmanifolds.* If G is a nilpotent group, then a sequence $g\colon \mathbb{N} \to G$ of the form $g(n) = a_1^{p_1(n)} a_2^{p_2(n)} \cdots a_k^{p_k(n)}$, where $a_i \in G$, and p_i are polynomials taking integer values at the integers, is called a *polynomial sequence in* G. If the maximum degree of the polynomials p_i is at most d we say that the *degree* of $g(n)$ is at most d. A *polynomial sequence on the nilmanifold* $X = G/\Gamma$ is a sequence of the form $(g(n)\Gamma)_{n\in\mathbb{N}}$ where $g\colon \mathbb{N} \to G$ is a polynomial sequence in G.

Theorem 2.7 (**Leibman** [33]). *Suppose that $X = G/\Gamma$ is a nilmanifold, with G connected and simply connected, and $(g(n))_{n\in\mathbb{N}}$ is a polynomial sequence in G. Let $Z = G/([G,G]\Gamma)$ and $\pi\colon X \to Z$ be the natural projection.*

Then the following statements are true:

(i) The sequence $(g(n)x)_{n\in\mathbb{N}}$ is equidistributed in a finite union of sub-nilmanifolds of X.

(ii) For every $x \in X$ the sequence $(g(n)x)_{n\in\mathbb{N}}$ is equidistributed in X if and only if the sequence $(g(n)\pi(x))_{n\in\mathbb{N}}$ is equidistributed in Z.

2.3.2. *Quantitative equidistribution on nilmanifolds.* We shall frequently use a quantitative version of Theorem 2.7 that was obtained in [23]. In order to state it we need to review some notions that were introduced in [23].

Given a nilmanifold $X = G/\Gamma$, the *horizontal torus* is defined to be the compact Abelian group $Z = G/([G,G]\Gamma)$. If X is connected, then Z is isomorphic to some finite dimensional torus $\mathbb{T}^l$. By $\pi\colon X \to H$ we denote the natural projection map. A *horizontal character* $\chi\colon G \to \mathbb{C}$ is a continuous homomorphism that satisfies $\chi(g\gamma) = \chi(g)$ for every $\gamma \in \Gamma$. Since every character annihilates G_2, every horizontal character factors through Z, and therefore can be thought of as a character of the horizontal torus. Since Z is identifiable with a finite dimensional torus $\mathbb{T}^l$ (we assume that X is connected), χ can also be thought of as a character of $\mathbb{T}^l$, in which case there exists a unique $\kappa \in \mathbb{Z}^l$ such that $\chi(t\mathbb{Z}^l) = e(\kappa \cdot t)$, where $\cdot$ denotes the inner product operation. We refer to κ as the *frequency* of χ and $\|\chi\| = |\kappa|$ as the *frequency magnitude* of χ.

Example 2.8. Let X be the Heisenberg nilmanifold (see Example 2.4). The map $\chi\colon G \to \mathbb{C}$ defined by $\chi\left(\begin{smallmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{smallmatrix}\right) = e(kx + ly)$, where $k, l \in \mathbb{Z}$, is a horizontal character of G. The map $\phi\left(\begin{smallmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{smallmatrix}\right) = (x\mathbb{Z}, y\mathbb{Z})$ induces an identification of the horizontal torus with $\mathbb{T}^2$. Under this identification, χ is mapped to the character $\tilde{\chi}(x\mathbb{Z}, y\mathbb{Z}) = e(kx + ly)$ of $\mathbb{T}^2$.

Suppose that $p\colon \mathbb{Z} \to \mathbb{R}$ is a polynomial sequence of degree k, then p can be uniquely expressed in the form $p(n) = \sum_{i=0}^{k} \binom{n}{i} \alpha_i$ where $\alpha_i \in \mathbb{R}$. We define

$$\|e(p(n))\|_{C^\infty[N]} = \max_{1\leq i\leq k} (N^i \|\alpha_i\|) \tag{5}$$

where $\|x\| = d(x, \mathbb{Z})$.

Given $N \in \mathbb{N}$, a finite sequence $(g(n)\Gamma)_{1\leq n\leq N}$ is said to be δ-*equidistributed*, if

$$\Big|\frac{1}{N}\sum_{n=1}^{N} F(g(n)\Gamma) - \int_X F\, dm_X\Big| \leq \delta \|F\|_{\mathrm{Lip}(X)}$$

for every Lipschitz function $F\colon X \to \mathbb{C}$, where

$$\|F\|_{\mathrm{Lip}(X)} = \|F\|_\infty + \sup_{x,y\in X, x\neq y} \frac{|F(x)-F(y)|}{d_X(x,y)}$$

for some appropriate metric d_X on X. We can now state the equidistribution result that we shall use. It is a direct consequence of Theorem 2.9 in [23] (we have suppressed some distracting quantitative details that will be of no use for us):

Theorem 2.9 (**Green & Tao** [23]). *Let $X = G/\Gamma$ be a nilmanifold with G connected and simply connected, and $d \in \mathbb{N}$.*

Then for every small enough $\delta > 0$ there exist $M = M_{X,d,\delta} \in \mathbb{R}$ with the following property: For every $N \in \mathbb{N}$, if $g\colon \mathbb{Z} \to G$ is a polynomial sequence of degree at most d such that the finite sequence $(g(n)\Gamma)_{1\leq n\leq N}$ is not δ-equidistributed, then for some non-trivial horizontal character χ with $\|\chi\| \leq M$ we have

$$\|\chi(g(n))\|_{C^\infty[N]} \leq M, \tag{6}$$

where χ is thought of as a character of the horizontal torus $Z = \mathbb{T}^l$ and $g(n)$ as a polynomial sequence in $\mathbb{T}^l$.

Example 2.10. It is instructive to interpret the previous result in some special case. Let $X = \mathbb{T}$ (with the standard metric), and suppose that the polynomial sequence on $\mathbb{T}$ is given by $p(n) = (n^d\alpha + q(n))\mathbb{Z}$ where $d \in \mathbb{N}$, $\alpha \in \mathbb{R}$, and $q \in \mathbb{Z}[x]$ with $\deg(q) \leq d-1$. In this case Theorem 2.9 reads as follows: There exists $M > 0$ such that for every $N \in \mathbb{N}$ and δ small enough, if the finite sequence $\big((n^d\alpha + q(n))\mathbb{Z}\big)_{1\leq n\leq N}$ is not δ-equidistributed in $\mathbb{T}$, then $\|k\alpha\| \leq M/N^d$ for some non-zero $k \in \mathbb{Z}$ with $|k| \leq M$.

3. A model equidistribution result

Before delving into the proof of the various equidistribution results on nilmanifolds we find it instructive to deal with a much simpler equidistribution problem on the circle. This model problem will motivate some of the ideas used later. We shall give a new proof for the following result:

Theorem 3.1 (**Boshernitzan** [8]). *Let $a \in \mathcal{H}$ have polynomial growth.*

Then the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$ if and only if for every $p \in \mathbb{Q}[t]$ we have $|a(t) - p(t)| \succ \log t$.

Our strategy will be to use the Taylor expansion of the function $a(t)$ to partition the range of the sequence $(a(n))_{n\in\mathbb{N}}$ into blocks that are approximately polynomial and then use classical results to estimate the corresponding exponential sums over these blocks. This argument can be adapted to the non-Abelian setup we are interested in, the main reason being that "Weyl type" sums involving polynomial block sequences of fixed degree on nilmanifolds can be effectively estimated using a rather sophisticated application of the van der Corput difference trick (this is done in [23]), and with a bit of care one can piece together these estimates to get usable results. The following simple example best illustrates our method:

Example 3.2. Suppose that $a(t) = t\log t$. We shall show that the sequence $(n\log n\,\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$. Using Lemma 3.3 below, it suffices to show that for every non-zero integer k we have

$$\lim_{N\to\infty} \mathbb{E}_{N<n\leq N+N^{\frac{3}{5}}}\, e(kn\log n) = 0.$$

For convenience we assume that $k = 1$.

Using the Taylor expansion of $a(t)$ around the point $x = N$ we see that for $n \in [1, N^{3/5}]$ we have

$$(N + n)\log(N + n) = N \log N + (\log N + 1)n + \frac{n^2}{2N} + o_{N\to\infty}(1). \tag{7}$$

(Notice that we keep track of the smallest order derivative that converges to zero and drop higher order derivatives.) Using (7) one gets

$$\mathbb{E}_{N<n\leq N+N^{\frac{3}{5}}} e(n \log n) = \mathbb{E}_{1\leq n\leq N^{\frac{3}{5}}} e\Big(N \log N + (\log N + 1)n + \frac{n^2}{2N}\Big) + o_{N\to\infty}(1). \tag{8}$$

Since $(N^{\frac{3}{5}})^2 \left\|\frac{1}{2N}\right\| \to \infty$, using Weyl's estimates (see e.g. [40]) we get that the averages in (8) converge to 0 as $N \to \infty$.

Lemma 3.3. *Let $(a(n))_{n\in\mathbb{N}}$ be a bounded sequence of complex numbers. Suppose that*

$$\lim_{N\to\infty} \big(\mathbb{E}_{N\leq n\leq N+l(N)} a(n)\big) = 0$$

for some positive function $l(t)$ with $l(t) \prec t$. Then

$$\lim_{N\to\infty} \mathbb{E}_{1\leq n\leq N} a(n) = 0.$$

Proof. We can cover the interval $[1, N]$ by a union I_N of non-overlapping intervals of the form $[k, k + l(k)]$. Since $l(t) \prec t$ and the sequence $(a(n))_{n\in\mathbb{N}}$ is bounded, we have that

$$\lim_{N\to\infty} \mathbb{E}_{1\leq n\leq N} a(n) = \lim_{N\to\infty} \mathbb{E}_{n\in I_N} a(n).$$

Using our assumption, one easily gets that the limit $\lim_{N\to\infty} \mathbb{E}_{n\in I_N} a(n)$ is zero, finishing the proof. □

A modification of the argument used in Example 3.2 gives the following more general result:

Lemma 3.4. *Suppose that for some $m \in \mathbb{N}$ the function $a \in C^{m+1}(\mathbb{R}_+)$ satisfies*

$$|a^{(m+1)}(t)| \text{ is decreasing}, \quad 1/t^m \prec a^{(m)}(t) \prec 1, \quad (a^{(m+1)}(t))^m \prec (a^{(m)}(t))^{m+1}.$$

Then the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$.

Proof. It suffices to show that for every non-zero integer k we have

$$\lim_{N\to\infty} \mathbb{E}_{1\leq n\leq N} e(ka(n)) = 0.$$

Since our assumptions are also satisfied for $ka(t)$ in place of $a(t)$ whenever $k \neq 0$, we can assume that $k = 1$.

By Lemma 3.3 it is enough to show that the averages

$$\mathbb{E}_{N\leq n\leq N+l(N)} e(a(n)) \tag{9}$$

converge to zero as $N \to \infty$ for some positive function $l(t)$ that satisfies $l(t) \prec t$.[2]

Using the Taylor expansion of $a(t)$ around the point $t = N$ we get

$$a(N + n) = a(N) + na'(N) + \cdots + \frac{n^m}{m!} a^{(m)}(N) + \frac{n^{m+1}}{(m + 1)!} a^{(m+1)}(\xi_n) \tag{10}$$

[2]The choice of $l(t)$ will depend on the function $a(t)$. For example, if $a(t) = t \log t$ we need to assume that $t^{1/2} \prec l(t) \prec t^{2/3}$, and if $a(t) = (\log t)^2$ we need to assume that $t/\log t \prec l(t) \prec t/\sqrt{\log t}$.

for some $\xi_n \in [N, N+n]$. Since $|a^{(m+1)}(t)|$ is decreasing we have $|a^{(m+1)}(\xi_n)| \leq |a^{(m+1)}(N)|$. It follows that if $l(t)$ also satisfies

$$(l(t))^{m+1} a^{(m+1)}(t) \prec 1,$$

then the averages in (9) are equal to

$$\mathbb{E}_{1\leq n\leq l(N)} e\Big(a(N) + na'(N) + \cdots + \frac{n^m}{m!} a^{(m)}(N)\Big) + o_{N\to\infty}(1).$$

Next, using Example 2.10 (or Weyl's estimates; see e.g. [40]) we get that the last averages converge to zero as $N \to \infty$ if

$$1 \prec (l(t))^m \left\| a^{(m)}(t) \right\| = (l(t))^m |a^{(m)}(t)|,$$

the last equality being valid for every large t since $|a^{(m)}(t)| \to 0$.

Summarizing, we have shown that the averages in (9) converge to zero when $N \to \infty$ as long as we can establish the existence of a function $l(t)$ satisfying the following conditions

$$l(t) \prec t \quad \text{and} \quad (l(t))^{m+1} |a^{(m+1)}(t)| \prec 1 \prec (l(t))^m |a^{(m)}(t)|. \tag{11}$$

Since by assumption $(a^{(m+1)}(t))^{\frac{m}{m+1}} \prec a^{(m)}(t)$ and $1/t^m \prec a^{(m)}(t)$, we can indeed find a function $l(t)$ that satisfies $\max\big((a^{(m+1)}(t))^{\frac{m}{m+1}}, 1/t^m\big) \prec 1/(l(t))^m \prec |a^{(m)}(t)|$, and so (11) holds. This completes the proof. □

The previous lemma applies to a wide variety of functions. For example the functions $(\log t)^2$, $t\log t$, $t^{3/2}$, $t^2\sqrt{2} + t^{1/2}$ satisfy the stated assumptions. In fact our next lemma shows that Lemma 3.4 comes rather close to establishing Theorem 3.1.

Lemma 3.5. *Let $a \in \mathcal{H}$ have polynomial growth and satisfy $|a(t) - p(t)| \succ \log t$ for every $p \in \mathbb{R}[t]$.*

Then the function $a(t)$ satisfies the assumptions of Lemma 3.4 for some $m \in \mathbb{N}$. As a consequence, the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$.

Proof. By Lemma 2.2 the function $a(t)$ has type k or k^+ for some non-negative integer k. We shall show that the assumptions of Lemma 3.4 are satisfied for $m = k+1$. We can assume that the function $a^{(k)}(t)$ is eventually positive, if this not the case we work with the function $-a(t)$.

Since $a(t) \prec t^{k+1}$, it follows from Corollary 2.3 that the functions $a^{(k+1)}(t)$ and $a^{(k+2)}(t)$ converge to zero. Furthermore, since both functions are elements of $\mathcal{H}$ the convergence is monotone.

We show that $a^{(k+1)}(t) \succ 1/t^{k+1}$. Suppose first that $a(t)$ has type k^+ for some positive integer k. By Corollary 2.3 we have

$$a^{(k+1)}(t) \succ a(t)/(t^{k+1}(\log t)^2) \gg 1/\big(t(\log t)^2\big) \succ 1/t^{k+1}.$$

Suppose now that $a(t)$ has type 0^+, in which case we shall show that $a'(t) \succ 1/t$. Arguing by contradiction, suppose that this is not the case. Since $a'(t)$ is eventually positive, we conclude that for large values of t we have $0 \leq a'(t) \leq c_1/t$ for some non-negative constant c_1. Integrating we get that for large values of t we have $0 \leq a(t) \leq c_1 \log t + c_2$ for some constants c_1, c_2, contradicting our assumption $|a(t)| \succ \log t$. Lastly, suppose that $a(t)$ has type k for some non-negative integer k. Since $a(t)$ stays away from polynomials, we conclude from Lemma 2.2 that $a(t) = p(t) + b(t)$, for some $p \in \mathbb{R}[t]$ of degree k, and some $b \in \mathcal{H}$ of type l^+ for some non-negative integer l with $l < k$. Arguing as before, we conclude that $b^{(k+1)}(t) \succ 1/t^{k+1}$. Since $a^{(k+1)}(t) = b^{(k+1)}(t)$, we get $a^{(k+1)}(t) \succ 1/t^{k+1}$.

It remains to show that $(a^{(k+2)}(t))^{k+1} \prec (a^{(k+1)}(t))^{k+2}$. By Lemma 2.1 we know that $a^{(k+2)}(t) \ll a^{(k+1)}(t)/t$. Using this, and the previously established estimate $a^{(k+1)}(t) \succ 1/t^{k+1}$, we get

$$(a^{(k+2)}(t))^{k+1} \ll (a^{(k+1)}(t))^{k+1}/t^{k+1} \prec (a^{(k+1)}(t))^{k+2}.$$

This completes the proof. □

We now complete the proof of Theorem 3.1

Proof of Theorem 3.1. We first prove the sufficiency of the conditions. Combining Lemma 3.4 and Lemma 3.5 we cover the case where $|a(t) - p(t)| \succ \log t$ for every $p \in \mathbb{R}[t]$. It remains to deal with the case where $a(t) = p(t) + e(t)$ for some $p \in \mathbb{R}[t]$ that has at least one non-constant coefficient irrational and $e(t) \ll \log t$. Since $e(n+1) - e(n) \to 0$ (this follows from the mean value theorem and the fact that $e'(t) \to 0$), we can write $\mathbb{N}$ as a union of non-overlapping intervals $(I_m)_{m\in\mathbb{N}}$ such that $|I_m| \to \infty$ and $\max_{n_1,n_2\in I_m} |e(n_1) - e(n_2)| \leq 1/m$. Combining this with the fact that the sequence $(p(n)\mathbb{Z})_{n\in\mathbb{N}}$ is well distributed in $\mathbb{T}$ (meaning $\lim_{N-M\to\infty} \mathbb{E}_{M\leq n\leq N} e(kp(n)) = 0$ for every non-zero $k \in \mathbb{Z}$), we deduce that the sequence $(a(n))_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$.

To prove the necessity of the conditions suppose that $1 \prec a(t) \ll \log t$; the general case can be easily reduced to this one. We shall show that the sequence $(a(n))_{n\in\mathbb{N}}$ cannot be equidistributed in $\mathbb{T}$. The key property we shall use is that $a(n+1) - a(n) \ll 1/n$. (This estimate is a consequence of the mean value theorem and the estimate $a'(t) \leq c/t$ for large enough t which can be proved as in Lemma 2.1.) For convenience we assume that $a(n+1) - a(n) < 1/n$ is satisfied for every $n \in \mathbb{N}$, and the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is increasing. The general case is similar. Arguing by contradiction, suppose that the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed. Let n_m be the first integer that satisfies $a(n_m) > m$. Since $a(n_m) < a(n) < a(n_m) + n/n_m$ and $a(n_m)$ is very close to an integer for large m, approximately all the integers in $[n_m, 3n_m/2]$ satisfy $\{a(n)\} \leq 1/2$. Furthermore, because of the equidistribution property, for large $m \in \mathbb{N}$, approximately half of the integers in $[1, n_m]$ satisfy $\{a(n)\} \leq 1/2$. Therefore, for large $m \in \mathbb{N}$, approximately two thirds of the integers in $[1, 3n_m/2]$ satisfy $\{a(n)\} \leq 1/2$, contradicting our equidistribution assumption. □

4. Single nil-orbits and Hardy sequences

In this section we are going to prove Theorems 1.1 and 1.2.

4.1. **A reduction.**

We start with some initial maneuvers that will allow us to reduce Theorem 1.2 to a more convenient statement.

First we give a result that enables us to translate distributional properties of sequences of the form $(b^{a(n)}x)_{n\in\mathbb{N}}$ to sequences of the form $(b^{[a(n)]}x)_{n\in\mathbb{N}}$.

Lemma 4.1. *Let $(a(n))_{n\in\mathbb{N}}$ be a sequence of real numbers such that for every nilmanifold $X = G/\Gamma$, with G connected and simply connected, and every $b \in G$, the sequence $(b^{a(n)}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in the nilmanifold $\overline{(b^s\Gamma)}_{s\in\mathbb{R}}$.*

Then for every nilmanifold $X = G/\Gamma$, every $b \in G$ and $x \in X$, the sequence $(b^{[a(n)]}x)_{n\in\mathbb{N}}$ is equidistributed in the nilmanifold $\overline{(b^n x)}_{n\in\mathbb{N}}$.

Proof. Let $X = G/\Gamma$ be a nilmanifold $b \in G$ and $x \in X$. We start with some reductions. By using the lifting argument of Section 2.2, we can assume that G is connected and simply connected. Furthermore, by changing the base point and using the formula in Section 2.2, we can assume that $x = \Gamma$.

Let X_b be the nilmanifold $\overline{(b^n\Gamma)}_{n\in\mathbb{N}}$ and m_{X_b} be the corresponding normalized Haar measure. It suffices to show that for every $F \in C(X)$ we have

$$\lim_{N\to\infty} \mathbb{E}_{1\leq n\leq N} F(b^{[a(n)]}\Gamma) = \int_{X_b} F\ dm_{X_b}. \tag{12}$$

So let $F \in C(X)$. To begin with, we use our assumption in the following case

$$\tilde{X} = \tilde{G}/\tilde{\Gamma} \ \text{ where } \ \tilde{G} = \mathbb{R}\times G, \ \tilde{\Gamma} = \mathbb{Z}\times\Gamma, \ \text{ and } \ \tilde{b} = (1, b).$$

(Notice that $\tilde{G}$ is connected and simply connected.) We conclude that for every $\tilde{H} \in C(\tilde{X})$

$$\lim_{N\to\infty} \mathbb{E}_{1\leq n\leq N} \tilde{H}(\tilde{b}^{a(n)}\tilde{\Gamma}) = \int_{\tilde{X}_{\tilde{b}}} \tilde{H}\ dm_{\tilde{X}_{\tilde{b}}}, \tag{13}$$

where $\tilde{X}_{\tilde{b}}$ is the nilmanifold $\overline{(s\mathbb{Z}, b^s\Gamma)}_{s\in\mathbb{R}}$, and $m_{\tilde{X}_{\tilde{b}}}$ is the corresponding normalized Haar measure.

Next we claim that (13) can be applied for the function $\tilde{F}\colon \tilde{X}\to\mathbb{C}$ defined by

$$\tilde{F}(t\mathbb{Z}, g\Gamma) = F(b^{-\{t\}}g\Gamma). \tag{14}$$

We caution the reader that the function $\tilde{F}$ may be discontinuous. The set of discontinuities of $\tilde{F}$ is a subset of the sub-nilmanifold $\{\mathbb{Z}\}\times X$. Near a point $(\mathbb{Z}, g\Gamma)$ of $\{\mathbb{Z}\}\times X$ the function $\tilde{F}$ comes close to the value $F(g\Gamma)$ or the value $F(b^{-1}g\Gamma)$. For $\delta > 0$ (and smaller than $1/2$) there exist functions $\tilde{F}_\delta \in C(\tilde{X})$ that agree with $\tilde{F}$ on $\tilde{X}_\delta = I_\delta \times X$, where $I_\delta = \{t\mathbb{Z}\colon \|t\| \geq \delta\}$, and are uniformly bounded by $2\|F\|_\infty$. Our assumption gives that the sequence $(a(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$. Since $\tilde{b}^{a(n)} = (a(n), b^{a(n)})$, we deduce that $\tilde{b}^{a(n)}\tilde{\Gamma} \in \tilde{X}_\delta$ for a set of $n \in \mathbb{N}$ with density $1 - 2\delta$. As a consequence,

$$\limsup_{N\to\infty} \mathbb{E}_{1\leq n\leq N} |\tilde{F}(\tilde{b}^{a(n)}\tilde{\Gamma}) - \tilde{F}_\delta(\tilde{b}^{a(n)}\tilde{\Gamma})| \leq 4\|F\|_\infty\,\delta. \tag{15}$$

By assumption, (13) holds when one uses the functions $\tilde{F}_\delta$ in place of the function $\tilde{H}$. Using these identities for every $\delta > 0$, and letting $\delta \to 0$, we get using (15) that (13) also holds for the discontinuous function $\tilde{F}$ defined in (14) (to get that $\int \tilde{F}_\delta\ dm_{\tilde{X}_{\tilde{b}}} \to \int \tilde{F}\ dm_{\tilde{X}_{\tilde{b}}}$ we use that $m_{\tilde{X}_{\tilde{b}}}(\{0\}\times X) = 0$, which holds since $\{0\}\times X$ is a proper sub-nilmanifold of $\tilde{X}_{\tilde{b}}$). This verifies our claim.

Applying (13) for the function $\tilde{F}$ defined in (14), and noticing that

$$\tilde{F}(\tilde{b}^{a(n)}\tilde{\Gamma}) = F(b^{-\{a(n)\}}b^{a(n)}\Gamma) = F(b^{[a(n)]}\Gamma),$$

we get

$$\lim_{n\to\infty} \mathbb{E}_{1\leq n\leq N} F(b^{[a(n)]}\Gamma) = \int_{\tilde{X}_{\tilde{b}}} \tilde{F}\ dm_{\tilde{X}_{\tilde{b}}} = \int_{\tilde{X}_{\tilde{b}}} F(b^{-\{s\}}g\Gamma)\ dm_{\tilde{X}_{\tilde{b}}}(s\mathbb{Z}, g\Gamma).$$

Since $b^{-\{s\}}b^s\Gamma = b^{[s]}\Gamma$, the map $(s\mathbb{Z}, g\Gamma) \to b^{-\{s\}}g\Gamma$ sends the nilmanifold $\tilde{X}_{\tilde{b}}$ onto the nilmanifold $X_b = \overline{(b^n\Gamma)}_{n\in\mathbb{N}}$. On X_b we define the measure m by letting

$$\int_{X_b} F\ dm = \int_{\tilde{X}_b} F(b^{-\{s\}}g\Gamma)\ dm_{\tilde{X}_{\tilde{b}}}(s\mathbb{Z}, g\Gamma)$$

for every $F \in C(X_b)$. We claim that $m = m_{X_b}$. Indeed, a quick computation shows that the measure m is invariant under left translation by b. As it is well known, any rotation b is uniquely ergodic on its orbit closure X_b, hence $m = m_{X_b}$. This establishes (12) and completes the proof. □

The previous lemma shows that part (ii) of Theorem 1.2 follows from part (i). It turns out that dealing with part (i) presents significant technical advantages (in fact we do not see how to establish part (ii) directly).

Next we show that in order to prove part (i) of Theorem 1.2 it suffices to establish the following result:

Proposition 4.2. *Let $a \in \mathcal{H}$ have polynomial growth and satisfy $|a(t) - cp(t)| \succ \log t$ for every $c \in \mathbb{R}$ and $p \in \mathbb{Z}[t]$. Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected, and suppose that $b \in G$ acts ergodically on X.*

Then the sequence $(b^{a(n)}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in X.

To carry out this reduction we shall use the following lemma:

Lemma 4.3. *Let $X = G/\Gamma$ be a nilmanifold with G connected and simply connected.*

Then for every $b \in G$ there exists $s_0 \in \mathbb{R}$ such that the element b^{s_0} acts ergodically on the nilmanifold $\overline{(b^s\Gamma)}_{s\in\mathbb{R}}$.

Proof. By Ratner's theorem (see Section 2.2), we have $\overline{(b^s\Gamma)}_{s\in\mathbb{R}} = H/\Delta$, where H is a connected and simply connected closed subgroup of G that contains all the elements b^s, $s \in \mathbb{R}$, and $\Delta = H \cap \Gamma$. By Theorem 2.7 it suffices to check that b^{s_0} acts ergodically on the horizontal torus $G/([G,G]|\Gamma)$, which we can assume to be $\mathbb{T}^k$ for some $k \in \mathbb{N}$. Equivalently, this amounts to showing that if $\beta\mathbb{Z}^k \in \mathbb{T}^k$, where $\beta \in \mathbb{R}^k$, then there exists $s_0 \in \mathbb{R}$ such that $\overline{(ns_0\beta\mathbb{Z}^k)}_{n\in\mathbb{N}} = \overline{(s\beta\mathbb{Z}^k)}_{s\in\mathbb{R}}$. One can check (we omit the routine details) that it suffices to choose s_0 such that the number $1/s_0$ is rationally independent of any non-zero integer combination of the coordinates of β. This completes the proof. □

Putting together Lemma 4.1 and Lemma 4.3 we get the advertised reduction:

Proposition 4.4. *In order to prove Theorem 1.2 it suffices to prove Proposition 4.2.*

Proof. Using Lemma 4.1, we see that part (ii) of Theorem 1.2 follows from part (i).

To establish part (i) we argue as follows. Let $b \in G$. By Lemma 4.3 there exists non-zero $s_0 \in \mathbb{R}$ such that the element b^{s_0} acts ergodically on the nilmanifold $\overline{(b^s\Gamma)}_{s\in\mathbb{R}}$. Using Proposition 4.2 for the element b^{s_0} and the function $a(s)/s_0$, we get that the sequence $(b^{a(n)}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in the nilmanifold $\overline{(b^s\Gamma)}_{s\in\mathbb{R}}$. □

We now turn our attention to the proof of Proposition 4.2.

4.2. **Proof of Proposition 4.2.** The following lemma is the key ingredient in the proof of Proposition 4.2:

Lemma 4.5. *Suppose that for some $k \in \mathbb{N}$ the function $a \in C^{k+1}(\mathbb{R}_+)$ satisfies*

$$|a^{(k+1)}(t)| \text{ is decreasing}, \quad 1/t^k \prec a^{(k)}(t) \prec 1, \quad (a^{(k+1)}(t))^k \prec (a^{(k)}(t))^{k+1}.$$

Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected, and suppose that $b \in G$ acts ergodically on X.

Then the sequence $(b^{a(n)}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in X.

Proof. Let $F \in C(X)$ with zero integral. We want to show that

$$\lim_{N\to\infty} \mathbb{E}_{1\le n\le N} F(b^{a(n)}\Gamma) = 0.$$

By Lemma 3.3 it suffices to show that the averages

$$\mathbb{E}_{N\leq n\leq N+l(N)}F(b^{a(n)}\Gamma) \tag{16}$$

converge to zero as $N\to\infty$ for some positive function $l(t)$ that satisfies $l(t)\prec t$.

Using the Taylor expansion of $a(t)$ around the point $x=N$ we have

$$a(N+n)=a(N)+na'(N)+\cdots+\frac{n^k}{k!}a^{(k)}(N)+\frac{n^{k+1}}{(k+1)!}a^{(k+1)}(\xi_n) \tag{17}$$

for some $\xi_n\in[N,N+n]$. Since $|a^{(k+1)}(t)|$ is decreasing we have $|a^{(k+1)}(\xi_n)|\leq|a^{(k+1)}(N)|$. It follows that if the function $l(t)$ satisfies

$$(l(t))^{k+1}a^{(k+1)}(t)\prec 1,$$

then the averages (16) are equal to

$$\mathbb{E}_{1\leq n\leq l(N)}F\Big(b^{p_N(n)}\Gamma\Big)+o_{N\to\infty}(1)$$

where

$$p_N(n)=a(N)+na'(N)+\cdots+\frac{n^k}{k!}a^{(k)}(N).$$

Our objective now is to show that for every $\delta>0$, for large values of N, the finite sequence $(b^{p_N(n)}\Gamma)_{1\leq n\leq l(N)}$ is δ-equidistributed in X. This would immediately imply that the averages in (16) converge to zero as $N\to\infty$.

So let $\delta>0$. Notice first that since

$$b^{p_N(n)}=b_{0,N}\,b_{1,N}^{n}\cdots b_{k,N}^{n^k},$$

where $b_{i,N}=b^{a^{(i)}(N)/i!}$ for $i=0,1,\ldots,k$, for every fixed $N\in\mathbb{N}$ the sequence $(b^{p_N(n)})_{n\in\mathbb{N}}$ is a polynomial sequence in G. Since $X=G/\Gamma$ and G is connected and simply connected, we can apply Theorem 2.9 (for δ small enough). We conclude that if the finite sequence $(b^{p_N(n)}\Gamma)_{1\leq n\leq l(N)}$ is not δ-equidistributed in X, then there exists a constant M (depending only on δ, X, and k), and a horizontal character χ with $\|\chi\|\leq M$ such that

$$\Big\|\chi(b^{p_N(n)})\Big\|_{C^\infty[l(N)]}\leq M. \tag{18}$$

Let $\pi(b)=(\beta_1\mathbb{Z},\ldots,\beta_s\mathbb{Z})$, where $\beta_i\in\mathbb{R}$, be the projection of b on the horizontal torus $\mathbb{T}^s$ (notice that s is bounded by the dimension of X). Since b acts ergodically on X the real numbers $1,\beta_1,\ldots,\beta_s$ must be rationally independent. For $t\in\mathbb{R}$ we have $\pi(b^t)=(t\tilde\beta_1\mathbb{Z},\ldots,t\tilde\beta_s\mathbb{Z})$ for some $\tilde\beta_i\in\mathbb{R}$ with $\tilde\beta_i\mathbb{Z}=\beta_i\mathbb{Z}$. As a consequence, we have

$$\chi(b^{p_N(n)})=e\Big(p_N(n)\sum_{i=1}^{s}l_i\tilde\beta_i\Big)$$

for some $l_i\in\mathbb{Z}$ with $|l_i|\leq M$. From this, the definition of $p_N(t)$, and the definition of $\|\cdot\|_{C^\infty[N]}$ (see (5)), we get that

$$\Big\|\chi(b^{p_N(n)})\Big\|_{C^\infty[l(N)]}\geq(l(N))^k\,\Big\|a^{(k)}(N)\beta\Big\|,$$

where β is a non-zero (we use the rational independence of the $\tilde{\beta}_i$'s here) real number that belongs to the finite set

$$B = \Big\{\frac{1}{k!}\sum_{i=1}^{s} l_i\tilde{\beta}_i \colon |l_i| \leq M\Big\}.$$

Combining this estimate with (18), and using that $\|a^{(k)}(N)\beta\| = |a^{(k)}(N)\beta|$ for large N (since by assumption $a^{(k)}(t) \to 0$), we get

$$(l(N))^k |a^{(k)}(N)\beta| \leq M \tag{19}$$

for some $\beta \in B$. It follows that if the function $l(t)$ satisfies

$$1 \prec (l(t))^k a^{(k)}(t),$$

then (19) fails for large N, and as a result the finite sequence $(b^{p_N(n)}\Gamma)_{1\leq n\leq l(N)}$ is δ-equidistributed in X for every large N.

Summarizing, we have shown that the averages (16) converge to 0 when $N \to \infty$, as long as we can find a positive function $l(t)$ that satisfies the following growth conditions

$$l(t) \prec t \ \text{ and } \ (l(t))^{k+1}a^{(k+1)}(t) \prec 1 \prec (l(t))^k a^{(k)}(t).$$

As in the proof of Lemma 3.4, one checks that the existence of such a function is guaranteed by our assumption, concluding the proof. $\square$

Notice that by Lemma 3.5 the previous result applies to every function $a \in \mathcal{H}$ that has polynomial growth and satisfies $|a(t) - p(t)| \succ \log t$ for every $p \in \mathbb{R}[t]$. In order to deal with the remaining cases of Proposition 4.2 we need one more lemma. Its proof follows the same strategy as in the proof of Lemma 4.5, so in order to avoid unnecessary repetition our argument will be rather sketchy.

Lemma 4.6. *Let $a \in \mathcal{H}$ satisfy $a(t) = p(t) + e(t)$, where $e(t) \ll \log t$, and $p \in \mathbb{R}[t]$ is not of the form $cq(t) + d$ with $c, d \in \mathbb{R}$ and $q \in \mathbb{Z}[t]$. Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected, and suppose that $b \in G$ acts ergodically on X.*

Then the sequence $(b^{a(n)}\Gamma)_{n\in\mathbb{N}}$ is equidistributed in X.

Remark. Our argument can easily be adapted to cover every function $e \in \mathcal{H}$ that satisfies $e(t) \prec t$, but the case treated suffices for our purposes.

Proof. Arguing as in Lemma 4.5, it suffices to show that for every $F \in C(X)$ with zero integral, the averages

$$\mathbb{E}_{N\leq n\leq N+\sqrt{N}} F(b^{a(n)}\Gamma) \tag{20}$$

converge to zero as $N \to \infty$.

Using Lemma 2.1 we conclude that the function $|e'(t)|$ is decreasing and $e'(t) \prec 1/t^{1-\varepsilon}$ for every $\varepsilon > 0$. Using the mean value theorem we conclude that for $n \in [1, \sqrt{N}]$ we have

$$e(N + n) = e(N) + o_{N\to\infty}(1)$$

and as a result the averages in (20) are equal to

$$\mathbb{E}_{1\leq n\leq\sqrt{N}} F\Big(b^{p(N+n)+e(N)}\Gamma\Big) + o_{N\to\infty}(1).$$

Hence, our proof will be complete if we show that for every $\delta > 0$, and every large N, the finite sequence $(b^{p(N+n)+e(N)}\Gamma)_{1\leq n\leq\sqrt{N}}$ is δ-equidistributed in X. Suppose that this is not the case. We are going to use Theorem 2.9 to derive a contradiction. The key property to be used

is that for every non-zero real number β the polynomial $\beta p(t)$ has at least one non-constant coefficient irrational. Arguing as in Lemma 4.5, we deduce that there exists a constant M (depending only on δ, X and the degree of p), and a finite set B of irrational numbers, such that for infinitely many positive integers N we have

$$\sqrt{N}\,\|\beta\| \leq M$$

for some $\beta \in B$. This is a contradiction and the proof is complete. □

Combining the last two lemmas it is now easy to prove Proposition 4.2.

Proof of Proposition 4.2 (Conclusion of proof of Theorem 1.2). Using Lemmas 3.5 and 4.5 we cover the case where $|a(t) - p(t)| \succ \log t$ for every $p \in \mathbb{R}[t]$. The remaining cases are covered by Lemma 4.6. □

Next we prove Theorem 1.1. It is a direct consequence of Theorem 1.2 and the following lemma:

Lemma 4.7. *Let $a \in \mathcal{H}$ satisfy $a(t) - p(t) \to 0$ for some $p \in \mathbb{R}[t]$.*

Then the sequences $(a(n))_{n\in\mathbb{N}}$ and $([a(n)])_{n\in\mathbb{N}}$ are pointwise good for nilsystems.

Proof. Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected, and $b \in G$. The sequence $(a(n))_{n\in\mathbb{N}}$ is pointwise good for nilsystems if and only if the same holds for the sequence $(p(n))_{n\in\mathbb{N}}$. Let $p(t) = c_0 + c_1 t + \cdots + c_k t^k$ for some non-negative integer k and $c_i \in \mathbb{R}$. Since $b^{p(n)} = b_0 \cdot b_1^n \cdot \ldots \cdot b_k^{n^k}$, where $b_i = b^{c_i}$, we have that $(b^{p(n)})_{n\in\mathbb{N}}$ is a polynomial sequence in G. It follows by Theorem 2.7 that the sequence $(p(n))_{n\in\mathbb{N}}$ is pointwise good for nilsystems.

Next we deal with the sequence $([a(n)])_{n\in\mathbb{N}}$. Suppose first that $p(t) - p(0) \in \mathbb{Q}[t]$. Then $p(t) = \frac{1}{r}\tilde{p}(t) + c$ for some $r \in \mathbb{N}$, $c \in \mathbb{R}$, and $\tilde{p} \in \mathbb{Z}[t]$ with $p(0) = 0$. For $i = 0, \ldots, r-1$ we have $[a(rn+i)] = q_i(n)$ for some $q_i \in \mathbb{Z}[t]$. Using this, the result follows from Theorem 2.7.

It remains to deal with the case where the polynomial p has an irrational non-constant coefficient. We are going to use a strategy similar to the one used in the proof of Lemma 4.1. Let

$$\tilde{X} = \tilde{G}/\tilde{\Gamma} \ \text{ where } \ \tilde{G} = \mathbb{R} \times G,\ \tilde{\Gamma} = \mathbb{Z} \times \Gamma, \ \text{ and } \ \tilde{b} = (1, b).$$

Given $F \in C(X)$ we define $\tilde{F} \colon \tilde{X} \to \mathbb{C}$ by

$$\tilde{F}(t\mathbb{Z}, g\Gamma) = F(b^{-\{t\}} g\Gamma).$$

(We caution the reader that $\tilde{F}$ may not be continuous). Notice that

$$F(b^{[a(n)]}\Gamma) = F(b^{-\{a(n)\}} b^{a(n)}\Gamma) = \tilde{F}(\tilde{b}^{a(n)}\tilde{\Gamma}),$$

and as a result it suffices to show that the averages

$$\mathbb{E}_{1\leq n\leq N} \tilde{F}(\tilde{b}^{a(n)}\tilde{\Gamma}) \tag{21}$$

converge as $N \to \infty$. We verify this as follows. For $\delta > 0$ (and smaller than $1/2$) there exist functions $\tilde{F}_\delta \in C(\tilde{X})$ that agree with $\tilde{F}$ on $\tilde{X}_\delta = I_\delta \times X$, where $I_\delta = \{t\mathbb{Z} \colon \|t\| \geq \delta\}$, and are uniformly bounded by $2\|F\|_\infty$. Since the polynomial p has a non-constant irrational coefficient, the sequence $(p(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$, and as a result $\tilde{b}^{p(n)}\tilde{\Gamma} \in \tilde{X}_\delta$ for a set of $n \in \mathbb{N}$ with density $1 - 2\delta$. It follows that

$$\limsup_{N\to\infty} \mathbb{E}_{1\leq n\leq N} |\tilde{F}(\tilde{b}^{a(n)}\tilde{\Gamma}) - \tilde{F}_\delta(\tilde{b}^{a(n)}\tilde{\Gamma})| \leq 4\|F\|_\infty \delta. \tag{22}$$

As shown in the first part of our proof, the sequence $(a(n))_{n\in\mathbb{N}}$ is pointwise good for nilsystems, hence the averages (21) converge when one uses the functions $\tilde{F}_\delta$ in place of the function $\tilde{F}$.

Using this and (22), we deduce that the averages in (21) form a Cauchy sequence, and hence they converge as $N \to \infty$. This proves that the sequence $([a(n)])_{n\in\mathbb{N}}$ is pointwise good for nilsystems and completes the proof. □

Proof of Theorem 1.1. The sufficiency of the conditions follows immediately from Theorem 1.2 and Lemma 4.7, with the exception of the case where $|a(t) - t/m| \ll \log t$ for some non-zero integer m. As noticed in [9] (proof of Theorem 3.3), this last case is easily reduced to the case $a(t) = t$. In this particular instance the result is well known (e.g. [33]).

The necessity of the conditions can be seen by working with rational rotations on the circle, for the details see [9]. □

5. Several nil-orbits and Hardy sequences

In this section we shall prove Theorem 1.3. A crucial part of our argument will be different than the one used to prove of Theorem 1.2, so we find it instructive to start with a model equidistribution problem that illustrates the key technical difference.

5.1. **A model equidistribution problem.** We shall give yet another proof of the following special case of Theorem 3.1:

"If $a \in \mathcal{H}$ satisfies $(t \log t) \prec a(t) \prec t^2$, then the sequence $(a(n))_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$."

We shall take the following fact for granted:

"If $a \in \mathcal{H}$ satisfies $\log t \prec a(t) \prec t$, then the sequence $(a(n))_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$."

So suppose that $a \in \mathcal{H}$ satisfies $(t \log t) \prec a(t) \prec t^2$. It suffices to show that for every non-zero $k \in \mathbb{Z}$ we have

$$\lim_{N\to\infty} \mathbb{E}_{1\le n\le N} e(ka(n)) = 0. \tag{23}$$

For convenience we assume that $k = 1$. For every fixed $R \in \mathbb{N}$ we have

$$\mathbb{E}_{1\le n\le RN} e(a(n)) = \mathbb{E}_{1\le n\le N}\big(\mathbb{E}_{1\le r\le R}\ e(a(Rn + r))\big) + o_{N\to\infty}(1). \tag{24}$$

For $n = 1, 2, \ldots$, we use the Taylor expansion of $a(t)$ around the point $t = Rn$. Since $a''(t) \to 0$ (by Lemma 2.1), we get for $r \in [1, R]$ that

$$a(Rn + r) = a(Rn) + ra'(Rn) + o_{n\to\infty;R}(1).$$

It follows that the averages in (24) are equal to

$$\mathbb{E}_{1\le n\le N} A_{R,n} + o_{N\to\infty;R}(1), \text{ where } A_{R,n} = \mathbb{E}_{1\le r\le R}\ e(a(Rn) + ra'(Rn)). \tag{25}$$

For fixed $\varepsilon > 0$ we split the averages $\mathbb{E}_{1\le n\le N}|A_{R,n}|$ as follows

$$\mathbb{E}_{1\le n\le N}(\mathbf{1}_{n\colon \|a'(Rn)\|\le\varepsilon} \cdot |A_{R,n}|) + \mathbb{E}_{1\le n\le N}(\mathbf{1}_{n\colon \|a'(Rn)\|>\varepsilon} \cdot |A_{R,n}|) = \Sigma_{1,R,N,\varepsilon} + \Sigma_{2,R,N,\varepsilon}.$$

We estimate $\Sigma_{1,R,N,\varepsilon}$. By Lemma 2.1 we have that $\log t \prec a'(Rt) \prec t$. It follows that the sequence $(a'(Rn)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$, and as a consequence

$$\frac{|1 \le n \le N\colon\ \|a'(Rn)\| \le \varepsilon|}{N} = 2\varepsilon + o_{N\to\infty;R}(1).$$

Therefore, $\Sigma_{1,R,N,\varepsilon} \le 2\varepsilon + o_{N\to\infty;R}(1)$.

We estimate $\Sigma_{2,R,N,\varepsilon}$. We have

$$|A_{R,n}| = |\mathbb{E}_{1\le r\le R}\ e(ra'(Rn))|.$$

We estimate the geometric series in the standard fashion; computing the sum and using the estimate $|\sin \pi t| \geq 2\,\|t\|$, we find that (whenever $a'(Rn)$ is not an integer)

$$|A_{R,n}| \leq \frac{1}{2R\,\|a'(Rn)\|}.$$

It follows that $\Sigma_{2,R,N,\varepsilon} \leq 1/(2R\varepsilon)$.

Combining the estimates for $\Sigma_{1,R,N,\varepsilon}$ and $\Sigma_{2,R,N,\varepsilon}$ we get

$$\mathbb{E}_{1\leq n\leq N}|A_{R,n}| \leq 2\varepsilon + \frac{1}{2R\varepsilon} + o_{N\to\infty;R}(1).$$

Letting first $N \to \infty$, then $R \to \infty$, and then $\varepsilon \to 0$, we get

$$\lim_{R\to\infty}\lim_{N\to\infty}\mathbb{E}_{1\leq n\leq N}|A_{R,n}| = 0.$$

As explained before, this implies (23) and completes the proof.

5.2. **A reduction.** As was the case with the proof of Theorem 1.2, we start with some initial maneuvers that enable us to reduce Theorem 1.3 to a more convenient statement. Since this step can be completed with straightforward modifications of the arguments used in Section 4.1, we omit the proofs.

First notice that in order to prove Theorem 1.3 we can assume that $X_1 = \cdots = X_\ell = X$. Indeed, consider the nilmanifold $\tilde{X} = X_1 \times \cdots \times X_\ell$. Then $\tilde{X} = \tilde{G}/\tilde{\Gamma}$, where $\tilde{G} = G_1 \times \cdots \times G_\ell$ is connected and simply connected, $\tilde{\Gamma} = \Gamma_1 \times \cdots \times \Gamma_\ell$ is a discrete cocompact subgroup of $\tilde{G}$, each b_i can be thought of as an element of $\tilde{G}$, and each x_i as an element of $\tilde{X}$.

Lemma 5.1. *Let $(a_1(n))_{n\in\mathbb{N}}, \dots, (a_\ell(n))_{n\in\mathbb{N}}$ be sequences of real numbers. Suppose that for every nilmanifold $X = G/\Gamma$, with G connected and simply connected, and every $b_1, \dots, b_\ell \in G$, the sequence*

$$(b_1^{a_1(n)}\Gamma, \dots, b_\ell^{a_\ell(n)}\Gamma)_{n\in\mathbb{N}}$$

is equidistributed in the nilmanifold $\overline{(b_1^s\Gamma)}_{s\in\mathbb{R}} \times \cdots \times \overline{(b_\ell^s\Gamma)}_{s\in\mathbb{R}}$.

Then for every nilmanifold $X = G/\Gamma$, every $b_1, \dots, b_\ell \in G$, and $x_1, \dots, x_\ell \in X$, the sequence

$$(b_1^{[a_1(n)]}x_1, \dots, b_\ell^{[a_\ell(n)]}x_\ell)_{n\in\mathbb{N}}$$

is equidistributed in the nilmanifold $\overline{(b_1^n x_1)}_{n\in\mathbb{N}} \times \cdots \times \overline{(b_\ell^n x_\ell)}_{n\in\mathbb{N}}$.

The previous lemma shows that part (ii) of Theorem 1.3 follows from part (i).

Lemma 5.2. *Let $X = G/\Gamma$ be a nilmanifold with G connected and simply connected.*

Then for every $b_1, \dots, b_\ell \in G$ there exists $s_0 \in \mathbb{R}$ such that for $i = 1, \dots, \ell$ the element $b_i^{s_0}$ acts ergodically on the nilmanifold $\overline{(b_i^s\Gamma)}_{s\in\mathbb{R}}$.

Using Lemmas 5.1 and 5.2, we see as in section Section 4.1, that Theorem 1.3 reduces to proving the following result:

Proposition 5.3. *Suppose that the functions $a_1(t), \dots, a_\ell(t)$ belong to the same Hardy field, have different growth rates, and satisfy $t^k \log t \prec a_i(t) \prec t^{k+1}$ for some $k = k_i \in \mathbb{N}$.*

Then for every nilmanifold $X = G/\Gamma$, with G connected and simply connected, and elements $b_1, \dots, b_\ell \in G$ acting ergodically on X, the sequence

$$(b_1^{a_1(n)}\Gamma, \dots, b_\ell^{a_\ell(n)}\Gamma)_{n\in\mathbb{N}}$$

is equidistributed in the nilmanifold X^ℓ.

5.3. **Proof of Proposition 5.3.** Since there is a key technical difference in the proofs of Proposition 5.3 and Proposition 4.2, we are going to give all the details. We are going to adapt the proof technique of the model equidistribution result of Section 5.1 to our particular non-Abelian setup.

Proof of Proposition 5.3. Let $F \in C(X^\ell)$ with zero integral. We want to show that

$$\lim_{N\to\infty} \mathbb{E}_{1\le n\le N} F(b_1^{a_1(n)}\Gamma, \ldots, b_\ell^{a_\ell(n)}\Gamma) = 0. \tag{26}$$

For every fixed $R \in \mathbb{N}$ we have

$$\mathbb{E}_{1\le n\le RN} F(b_1^{a_1(n)}\Gamma, \ldots, b_\ell^{a_\ell(n)}\Gamma) = \mathbb{E}_{1\le n\le N}\big(\mathbb{E}_{1\le r\le R}\, F(b_1^{a_1(nR+r)}\Gamma, \ldots, b_\ell^{a_\ell(nR+r)}\Gamma)\big) + o_{N\to\infty;R}(1). \tag{27}$$

For $n = 1, 2, \ldots$, we use the Taylor expansion of the functions $a_i(t)$ around the point $t = Rn$. Since $t^{k_i}\log t \prec a_i(t) \prec t^{k_i+1}$ for some $k_i \in \mathbb{N}$, Lemma 2.1 gives that $a_i^{(k_i+1)}(t) \to 0$. Hence, for $r \in [1, R]$ we have that

$$a_i(Rn + r) = p_{i,R,n}(r) + o_{n\to\infty;R}(1),$$

where

$$p_{i,R,n}(r) = a_i(Rn) + r a_i'(Rn) + \cdots + \frac{r^{k_i}}{k_i!} a_i^{(k_i)}(Rn). \tag{28}$$

It follows that the averages in (27) are equal to

$$\mathbb{E}_{1\le n\le N} A_{R,n} + o_{N\to\infty;R}(1), \text{ where } A_{R,n} = \mathbb{E}_{1\le r\le R}\, F(b_1^{p_{1,R,n}(r)}\Gamma, \ldots, b_\ell^{p_{\ell,R,n}(r)}\Gamma). \tag{29}$$

Our objective now is to show that for every $\delta > 0$, for all large values of R, the finite sequence $(b_1^{p_{1,R,n}(r)}, \ldots, b_\ell^{p_{\ell,R,n}(r)})_{1\le r\le R}$ is δ-equidistributed in X^ℓ for most values of n. This will enable us to show that the averages in (29) converge to zero as $N \to \infty$.

So let $\delta > 0$. As in the proof of Proposition 4.2 we verify that for fixed $R, n \in \mathbb{N}$ the sequence $(b_1^{p_{1,R,n}(r)}, \ldots, b_\ell^{p_{\ell,R,n}(r)})_{r\in\mathbb{N}}$ is a polynomial sequence in G^k. Since $X^\ell = G^\ell/\Gamma^\ell$, and G^ℓ is connected and simply connected, we can apply Theorem 2.9 (for small δ). We get that if the finite sequence $(b_1^{p_{1,R,n}(r)}\Gamma, \ldots, b_\ell^{p_{\ell,R,n}(r)}\Gamma)_{1\le r\le R}$ is not δ-equidistributed in X^ℓ, then there exists a constant M (depending only on δ, X, and the k_i's), and a non-trivial horizontal character χ of X^ℓ, with $\|\chi\| \le M$, and such that

$$\left\|\chi(b_1^{p_{1,R,n}(r)}, \ldots, b_\ell^{p_{\ell,R,n}(r)})\right\|_{C^\infty[R]} \le M. \tag{30}$$

For $i = 1, \ldots, \ell$, let $\pi(b_i) = (\beta_{i,1}\mathbb{Z}, \ldots, \beta_{i,s}\mathbb{Z})$, where $\beta_{i,j} \in \mathbb{R}$, be the projection of b_i on the horizontal torus $\mathbb{T}^s$ of X (notice that s is bounded by the dimension of X). Since each b_i acts ergodically on X, the set of real numbers $\{1, \beta_{i,1}, \ldots, \beta_{i,s}\}$ is rationally independent for $i = 1, \ldots, \ell$. For $t \in \mathbb{R}$ we have $\pi(b_i^t) = (t\tilde{\beta}_{i,1}\mathbb{Z}, \ldots, t\tilde{\beta}_{i,s}\mathbb{Z})$ for some $\tilde{\beta}_{i,j} \in \mathbb{R}$ with $\tilde{\beta}_{i,j}\mathbb{Z} = \beta_{i,j}\mathbb{Z}$. As a consequence

$$\chi(b_1^{p_{1,R,n}(r)}, \ldots, b_\ell^{p_{\ell,R,n}(r)}) = e\Big(\sum_{i=1}^{s} \big(p_{i,R,n}(n) \sum_{j=1}^{s} l_{i,j}\tilde{\beta}_{i,j}\big)\Big) \tag{31}$$

for some $l_{i,j} \in \mathbb{Z}$ with $|l_{i,j}| \le M$. Let $k_{\min} = \min\{k_1, \ldots, k_\ell\}$ and $k_{\max} = \max\{k_1, \ldots, k_\ell\}$. It follows from (28), (31), and the definition of $\|\cdot\|_{C^\infty[R]}$ (see (5)), that there exists $k \in \mathbb{N}$ with

$k_{\min} \le k \le k_{\max}$ such that

$$\Big\|\chi(b_1^{p_{1,R,n}(r)}, \dots, b_\ell^{p_{\ell,R,n}(r)})\Big\|_{C^\infty[R]} \ge R^k \Big\| \sum_{i\in I} a_i^{(k)}(Rn)\beta_i \Big\|,$$

where the sum ranges over those $i \in \{1, \dots, \ell\}$ that satisfy $k_i = k$, and the β_i's are real numbers, not all of them zero (we used here that χ is non-trivial and the rational independence of the $\tilde\beta_{i,j}$'s), that belong to the finite set

$$B = \bigcup_{i=1}^{\ell} \Big\{ \frac{1}{k!} \sum_{j=1}^{s} l_{i,j}\tilde\beta_{i,j} \colon |l_{i,j}| \le M \Big\}.$$

Combining this estimate with (30) gives

$$\Big\| \sum_{i\in I} a_i^{(k)}(Rn)\beta_i \Big\| \le \frac{M}{R^k} \tag{32}$$

for some $\beta_i \in B$.

We are now ready to estimate the average $\mathbb{E}_{1\le n\le N}|A_{R,n}|$. Given $\varepsilon > 0$ we split it as follows
$\mathbb{E}_{1\le n\le N}|A_{R,n}| = \mathbb{E}_{1\le n\le N}(\mathbf{1}_{S_{1,R,\varepsilon}}(n)\cdot|A_{R,n}|) + \mathbb{E}_{1\le n\le N}(\mathbf{1}_{S_{2,R,\varepsilon}}(n)\cdot|A_{R,n}|) = \Sigma_{1,R,N,\varepsilon} + \Sigma_{2,R,N,\varepsilon}$,
where

$$S_{1,R,\varepsilon} = \Big\{ n\in\mathbb{N}\colon \Big\|\sum_{i\in I} a_i^{(k)}(Rx)\beta_i\Big\| \le \varepsilon \text{ for some } \beta_i\in B, \text{ not all of them } 0\Big\},\ S_{2,R,\varepsilon} = \mathbb{N}\setminus S_{1,R,\varepsilon}.$$

We estimate $\Sigma_{1,R,N,\varepsilon}$. Using Lemma 2.1 and our assumptions, we conclude that $\log t \prec a_i^{(k)}(t) \prec t$ for $i\in I$. Furthermore, since the functions $a_i(t)$ for $i \in I$ have different growth rates and belong to the same Hardy field, we deduce that the functions $a_i^{(k)}(t)$ for $i\in I$ have different growth rates. It follows that

$$\log t \prec b_R(t) = \sum_{i\in I} a_i^{(k)}(Rt)\beta_i \prec t.$$

Since $b_R \in \mathcal{H}$ and $\log t \prec b_R(t) \prec t$, we get (e.g. using Theorem 3.1) that for every $R\in\mathbb{N}$ the sequence $(b_R(n)\mathbb{Z})_{n\in\mathbb{N}}$ is equidistributed in $\mathbb{T}$. Hence,

$$\frac{|1\le n\le N\colon\ \|b_R(n)\|\le\varepsilon|}{N} = 2\varepsilon + o_{N\to\infty;R}(1).$$

It follows that

$$|\Sigma_{1,R,N,\varepsilon}| \le 2\,\|F\|_\infty\,\varepsilon + o_{N\to\infty;R}(1).$$

We estimate $\Sigma_{2,R,N,\varepsilon}$. Notice that for $n\in S_{2,R,\varepsilon}$ we have $\Big\|\sum_{i\in I} a_i^{(k)}(Rn)\beta_i\Big\| \ge \varepsilon$. As a result, if R is large enough, then (32) fails, and as a consequence the finite sequence $(b_1^{p_{1,R,n}(r)}x_1, \dots, b_\ell^{p_{\ell,R,n}(r)}x_\ell)_{1\le r\le R}$ is δ-equidistributed in X. Hence, if R is large enough, then $|A_{R,n}| \le \delta$ for every $n\in S_{2,R,\varepsilon}$. Therefore, for every $N\in\mathbb{N}$ we have

$$|\Sigma_{2,R,N,\varepsilon}| \le \delta + o_{R\to\infty}(1).$$

Putting the previous estimates together we find

$$\mathbb{E}_{1\le n\le N}|A_{R,n}| \le 2\,\|F\|_\infty\,\varepsilon + \delta + o_{N\to\infty;R}(1) + o_{R\to\infty}(1).$$

Letting $N\to\infty$, then $R\to\infty$, and then $\varepsilon,\delta\to 0$, we deduce that

$$\lim_{R\to\infty}\lim_{N\to\infty} \mathbb{E}_{1\le n\le N}|A_{R,n}| = 0,$$

or equivalently that

$$\lim_{R\to\infty}\lim_{N\to\infty}\mathbb{E}_{1\leq n\leq N}\Big|\mathbb{E}_{1\leq r\leq R}F(b^{[a_1(Rn+r)]}\Gamma,\ldots,b^{[a_\ell(Rn+r)]}\Gamma)\Big|=0.$$

Combining this with (27) gives (26), completing the proof. □

6. Random sequences of sub-exponential growth

In this section we shall prove Theorem 1.4. In what follows, when we introduce a nilpotent Lie group G or a nilmanifold X, we assume that it comes equipped with a Mal'cev basis and the corresponding (right invariant) metric d_G or d_X that was introduced in [23]. When there is no danger of confusion we are going to denote d_G or d_X with d. We denote by B_M the ball in G of radius M, that is $B_M=\{g\in G\colon d(g,\mathrm{id}_G)\leq M\}$.

6.1. **A reduction.**

We start with some initial maneuvers that will allow us to reduce Theorem 1.4 to a more convenient statement.

We remind the reader of our setup. We are given a sequence $(X_n(\omega))_{n\in\mathbb{N}}$ of $0-1$ valued independent random variables with $P(\{\omega\in\Omega\colon X_n(\omega)=1\})=\sigma_n$, where $(\sigma_n)_{n\in\mathbb{N}}$ is a decreasing sequence of real numbers satisfying $\lim_{n\to\infty}n\sigma_n=\infty$. Our objective is to show that almost surely the averages

$$\frac{1}{N}\sum_{n=1}^{N}F(b^{a_n(\omega)}\Gamma) \tag{33}$$

converge as $N\to\infty$ for every nilmanifold $X=G/\Gamma$, function $F\in C(X)$, and element $b\in G$.

We caution the reader that the set of probability 1 for which the averages (33) converge has to be independent of the nilmanifold $X=G/\Gamma$, the function $F\in C(X)$, and the element $b\in G$. On the other hand, since up to isomorphism there exist countably many nilmanifolds X (see for example [12]), and since the space $C(X)$ is separable, it suffices to prove that for every fixed nilmanifold $X=G/\Gamma$ and $F\in C(X)$ the averages (33) converge almost surely for every $b\in G$. Furthermore, since G is a countable union of balls, it suffices to verify the previous statement with B_M in place of G for every $M>0$.

Next notice that instead of working with the averages (33), it suffices to work with the averages

$$\frac{1}{A(N,\omega)}\sum_{n=1}^{N}X_n(\omega)\,F(b^n\Gamma)$$

where $A(N,\omega)=|\{n\in\{1,\ldots,N\}\colon X_n(\omega)=1\}|$. Since the expectation of X_n is σ_n, by the strong law of large numbers we almost surely have that $A(N,\omega)/w(N)\to 1$, where $w(N)=\sum_{n=1}^{N}\sigma_n$. It therefore suffices to work with the averages

$$\frac{1}{w(N)}\sum_{n=1}^{N}X_n(\omega)\,F(b^n\Gamma).$$

We shall establish convergence of these averages by comparing them with the averages

$$\frac{1}{w(N)}\sum_{n=1}^{N}\sigma_n\,F(b^n\Gamma).$$

Notice that these last averages can be compared with the averages

$$\frac{1}{N}\sum_{n=1}^{N} F(b^n\Gamma)$$

which, as we have mentioned repeatedly before, are known to be convergent.

Up to this point we have reduced matters to showing that for every nilmanifold $X = G/\Gamma$, $F \in C(X)$, and $M > 0$, we almost surely have

$$\lim_{N\to\infty} \frac{1}{w(N)} \sum_{n=1}^{N} (X_n(\omega) - \sigma_n)\, F(b^n\Gamma) = 0 \tag{34}$$

for every $b \in B_M$.

Next we show that we can impose a few extra assumptions on the nilmanifold X, and the function $F \in C(X)$. Using the lifting argument of Section 2.2 we see that every sequence $(F(b^n\Gamma))_{n\in\mathbb{N}}$ can be represented in the form $(\tilde{F}(\tilde{b}^n\tilde{\Gamma}))_{n\in\mathbb{N}}$ for some nilmanifold $\tilde{X} = \tilde{G}/\tilde{\Gamma}$, with $\tilde{G}$ connected and simply connected, $\tilde{F} \in C(\tilde{X})$, and $\tilde{b} \in \tilde{G}$. Therefore, when proving (34) we can assume that the nilmanifold X has the form G/Γ, where the group G is connected and simply connected. Furthermore, since the set $\text{Lip}(X)$, of Lipschitz functions $F\colon X \to \mathbb{C}$, is dense in $C(X)$ in the uniform topology, an easy approximation argument shows that it suffices to prove (34) for $F \in \text{Lip}(X)$.

Summarizing, we have reduced Theorem 1.4 to proving:

Theorem 6.1. *Let $(X_n(\omega))_{n\in\mathbb{N}}$ be a sequence of $0-1$ valued independent random variables with $P(\{\omega \in \Omega\colon X_n(\omega) = 1\}) = \sigma_n$, where $(\sigma_n)_{n\in\mathbb{N}}$ is a decreasing sequence of real numbers satisfying $\lim_{n\to\infty} n\sigma_n = \infty$. Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected, $F \in Lip(X)$, and $M > 0$.*

Then almost surely we have

$$\lim_{N\to\infty} \max_{b\in B_M} \left| \frac{1}{w(N)} \sum_{n=1}^{N} (X_n(\omega) - \sigma_n)\, F(b^n\Gamma) \right| = 0$$

where $w(N) = \sum_{n=1}^{N} \sigma_n$.

Remark. We shall not use the fact that the convergence to zero is uniform; only the independence of the set of full measure on the set B_M will be used.

To prove Theorem 6.1 we are going to extend an argument used by Bourgain in [10] (where the case $X = \mathbb{T}$ was covered). A more detailed version of this argument can be found in [38]. Since several steps of [38] carry over verbatim to our case we are only going to spell out the details of the genuinely new steps.

6.2. **A key ingredient.** In this subsection we shall prove the following key result:

Proposition 6.2. *Let $X = G/\Gamma$ be a nilmanifold, with G connected and simply connected, and M be a positive real number.*

Then there exists $k = k(G, M) \in \mathbb{N}$ with the following property: for every $N \in \mathbb{N}$, there exists $B_{N,M} \subset B_M$ with $|B_{N,M}| = N^k$, and such that for every $F \in Lip(X)$ with $\|F\|_{Lip(X)} \le 1$, and sequence of real numbers $(c_n)_{n\in\mathbb{N}}$ with norm bounded by 1 , we have

$$\max_{b\in B_M} \Big| \sum_{n=1}^{N} c_n\, F(b^n\Gamma) \Big| = \max_{b\in B_{N,M}} \Big| \sum_{n=1}^{N} c_n\, F(b^n\Gamma) \Big| + o_{N\to\infty}(1). \tag{35}$$

The proof of this result ultimately relies on the fact that multiplication on a nilpotent Lie group is given by polynomial mappings. To make this precise we shall use a convenient coordinate system, the proof of its existence can be found in [23] (for example).

For every connected and simply connected Lie group G there exist a non-negative integer m (we call m the *dimension* of G) and a continuous isomorphism ϕ from $(G, \cdot)$ to $(\mathbb{R}^m, \cdot)$ with multiplication defined as follows: If $u = (u_1, \ldots, u_m)$ and $v = (v_1, \ldots, v_m)$, then for $i = 1, \ldots, m$ the i-th coordinate of $u \cdot v$ has the form

$$u_i + v_i + P_i(u_1, \ldots, u_{i-1}, v_1, \ldots, v_{i-1})$$

where $P_i \colon \mathbb{R}^{i-1} \times \mathbb{R}^{i-1} \to \mathbb{R}$ is a polynomial of degree at most i. It follows that the i-th coordinate of u^n has the form

$$nu_i + Q_i(u_1, \ldots, u_{i-1}, n)$$

where $Q_i \colon \mathbb{R}^{i-1} \times \mathbb{R} \to \mathbb{R}$ is a polynomial.

We shall use the following result (Lemma A.4 in [23]):

Lemma 6.3 (**Green & Tao** [23])**.** *Let G be a connected and simply connected nilpotent Lie group of dimension m.*

Then there exists $k = k(G) \in \mathbb{N}$ such that for every $K > 1$ we have

$$K^{-k}|u - v| \leq d(g, h) \leq K^k|u - v|,$$

for every $g, h \in G$, and $u = \phi(g), v = \phi(h) \in \mathbb{R}^m$ that satisfy $|u|$, $|v| \leq K$, where $|\cdot|$ denotes the sup-norm in $\mathbb{R}^m$.

Using this, we are going to show:

Lemma 6.4. *Let G be a connected and simply connected nilpotent Lie group.*

Then there exists $k = k(G) \in \mathbb{N}$ such that for every $M > 0$ we have

$$d(g^n, h^n) \ll_{G,M} n^k d(g, h)$$

for every $n \in \mathbb{N}$ and $g, h \in B_M$.

Proof. We first establish the corresponding estimate in "coordinates". Suppose that the dimension of G is m. Let $\phi(g) = u = (u_1, \ldots, u_m)$ and $\phi(h) = v = (v_1, \ldots, v_m)$ satisfy $|u|, |v| \leq K$. Using the multiplication formula in local coordinates we deduce that

$$|(u^n)_i - (v^n)_i| \leq \sum_{j=1}^{i} |(u_j - v_j)|\, |R_j(u_1, \ldots, u_{j-1}, v_1, \ldots, v_{j-1}, n)|$$

for some polynomials $R_j \colon \mathbb{R}^{j-1} \times \mathbb{R}^{j-1} \times \mathbb{R} \to \mathbb{R}$ of degree of degree depending only on G. If we consider R_j as a polynomial of a single variable n, then its coefficients depend polynomially on the parameters u_i, v_i (which are bounded by K) and the structure constants of the Mal'cev basis of G. Hence, $|R_j(u, v, n)| \ll_{G,K} n^{l_j}$ for some $l_j = l_j(G) \in \mathbb{N}$. It follows that

$$|(u^n)_i - (v^n)_i| \ll_{G,K} n^{k_1} \sum_{j=1}^{i} |(u_j - v_j)|$$

for some $k_1 = k_1(G)$. As a consequence

$$|u^n - v^n| \ll_{G,K} n^{k_1} |u - v| \tag{36}$$

for some $k_1 = k_1(G)$.

To finish the proof, we use (36) to deduce an analogous estimate for elements of G with the metric d. We argue as follows. First, using Lemma 6.3 we conclude that if $g \in B_M$, then $|u| \ll_{G,M} 1$. As a result, (36) gives that

$$|u^n| \ll_{G,M} n^{k_1} \tag{37}$$

for every $g \in B_M$. Next, notice that by Lemma 6.3 there exists $k_2 = k_2(G) \in \mathbb{N}$ such that for every $K > 1$ we have

$$K^{-k_2}|u - v| \leq d(g, h) \leq K^{k_2}|u - v| \tag{38}$$

for every $g, h \in G$ that satisfy $|u|$, $|v| \leq K$ (remember that $u = \phi(g), v = \phi(h) \in \mathbb{R}^m$). Combining the estimates (36), (37), and (38) we get

$$d(g^n, h^n) \ll_{G,M} n^{k_1 k_2}|u^n - v^n| \ll_{G,M} n^{k_1+k_1k_2}|u - v| \ll_{G,M} n^{k_1+k_1k_2} d(g, h).$$

This establishes the advertised estimate with $k = k_1 + k_1 k_2$. □

Proof of Proposition 6.2. By Lemma 6.4 we get that there exists $k_1 = k_1(G)$ such that

$$d_G(g^n, h^n) \ll_{G,M} n^{k_1} d_G(g, h) \tag{39}$$

for every $g, h \in B_M$. For every $K \in \mathbb{N}$ there exist K^m points that form an $1/K$-net for the set $[0, 1)^m$ with the sup-norm. Combining this with Lemma 6.3 we get that there exists $k_2 = k_2(G)$ with the following property: for every $N \in \mathbb{N}$ there exists an $1/N^{k_1+2}$ net of B_M consisting of N^{k_2} points.

Let $B_{N,M}$ be any such $1/N^{k_1+2}$-net of B_M. By construction, $|B_{N,M}| = N^{k_2}$ for some k_2 that depends only on G. Furthermore, for every $b \in B_M$ there exists $b_N \in B_{N,M}$ such that $d_G(b, b_N) \leq 1/N^{k_1+2}$. It follows from (39) that

$$\max_{1\leq n\leq N} d_G(b^n, b_N^n) \ll_{G,M} N^{k_1} d_G(b, b_N) \leq 1/N^2.$$

Therefore,

$$\max_{1\leq n\leq N} d_X(b^n\Gamma, b_N^n\Gamma) \ll_{G,M} 1/N^2.$$

Using this, we deduce (35) (with $o_{N\to\infty}(1) = \|c_n\|_\infty \|F\|_{\mathrm{Lip}(X)} /N \leq 1/N$), completing the proof. □

6.3. **Proof of Theorem 6.1.** We give a sketch of the proof of Theorem 6.1. The missing details can be extracted from [38].

Without loss of generality we can assume that $\|F\|_{\mathrm{Lip}(X)} \leq 1$.

From Proposition 6.2 we conclude that there exists a $k = k(G, M) \in \mathbb{N}$ and a subset $B_{N,M}$ of B_M with $|B_{N,M}| = N^k$ such that

$$\max_{b\in B_M} \left|\frac{1}{w(N)} \sum_{n=1}^{N} (X_n(\omega) - \sigma_n) F(b^n\Gamma)\right| = \max_{b\in B_{N,M}} \left|\frac{1}{w(N)} \sum_{n=1}^{N} (X_n(\omega) - \sigma_n) F(b^n\Gamma)\right| + o_{N\to\infty}(1). \tag{40}$$

Since the cardinality of $B_{N,M}$ is a power of N that depends only on G and M, it follows that $|B_{N,M}|^{1/\log N}$ is bounded by some constant that depends only on G and M. Hence,

(41)

$$\left\|\max_{b\in B_{N,M}}\left|\frac{1}{w(N)}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\,F(b^n\Gamma)\right|\right\|_{L^{\log N}(\Omega)} \ll_{G,M} \max_{b\in B_{N,M}}\left\|\frac{1}{w(N)}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\,F(b^n\Gamma)\right\|_{L^{\log N}(\Omega)}.$$

Furthermore, arguing exactly as in [38] (pages 40-41), it can be shown that for every sequence of complex numbers $(c_n)_{n\in\mathbb{N}}$ with $\|c_n\|_\infty \leq 1$ one has

$$\left\|\frac{1}{w(N)}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\,c_n\right\|_{L^{\log N}(\Omega)} \ll \sqrt{\frac{\log N}{w(N)}}. \tag{42}$$

Combining (40), (41), and (42) (with $c_n = F(b^n\Gamma)$), gives

$$\left\|\max_{b\in B_M}\left|\frac{1}{w(N)}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\,F(b^n\Gamma)\right|\right\|_{L^{\log N}(\Omega)} \ll_{G,M} \sqrt{\frac{\log N}{w(N)}}+o_{N\to\infty}(1). \tag{43}$$

Next we make use of the following simple lemma:

Lemma 6.5. *Let $(Y_k)_{k\in\mathbb{N}}$ be a sequence of bounded, complex-valued random variables on a probability space (Ω,Σ,P).*

Then almost surely we have

$$\limsup_{k\to\infty}\frac{|Y_k(\omega)|}{\|Y_k\|_{L^{\log k}(\Omega)}} \leq e$$

where e is the Euler number.

Combining this lemma with (43), we conclude that for every nilmanifold $X = G/\Gamma$, with G connected and simply connected, and $F \in \mathrm{Lip}(X)$ with $\|F\|_{\mathrm{Lip}(X)} \leq 1$, there exists a set $\Omega_{F,G,M}$ of probability 1, such that for every $\omega \in \Omega_{F,G,M}$ we have

$$\max_{b\in B_M}\left|\frac{1}{w(N)}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\,F(b^n\Gamma)\right| \ll_{\omega,G,M} \sqrt{\frac{\log N}{w(N)}}+o_{N\to\infty}(1).$$

Since by assumption $\log N/w(N) \to 0$, we get that for every nilmanifold X, $F \in \mathrm{Lip}(X)$, and $M > 0$, we almost surely have

$$\lim_{N\to\infty}\max_{b\in B_M}\left|\frac{1}{w(N)}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\,F(b^n\Gamma)\right| = 0.$$

This completes the proof of Theorem 6.1, and finishes the proof of Theorem 1.4.

References

[1] L. Auslander, L. Green, F. Hahn. Flows on homogeneous spaces. With the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg, *Annals of Mathematics Studies*, **53**, Princeton University Press, Princeton, N.J. (1963).

[2] V. Bergelson, B. Host, B. Kra, with an appendix by I. Ruzsa. Multiple recurrence and nilsequences. *Inventiones Math.* **160** (2005), 261–303.

[3] V. Bergelson, A. Leibman. Distribution of values of bounded generalised polynomials. *Acta Mathematica* **198** (2007), 155–230.
[4] V. Bergelson, A. Leibman, E. Lesigne. Intersective polynomials and the polynomial Szemerédi theorem. *Adv. Math.* **219** (2008), no. 1, 369–388.
[5] M. Boshernitzan. An extension of Hardy's class L of "Orders of Infinity". *J. Analyse Math.* **39** (1981), 235–255.
[6] M. Boshernitzan. New "Orders of Infinity". *J. Analyse Math.* **41** (1982), 130–167.
[7] M. Boshernitzan. Homogeneously distributed sequences and Poincaré sequences of integers of sublacunary growth. *Monatsh. Math.* **96** (1983), no. 3, 173–181.
[8] M. Boshernitzan. Uniform distribution and Hardy fields. *J. Analyse Math.* **62** (1994), 225–240.
[9] M. Boshernitzan, G. Kolesnik, A. Quas, M. Wierdl. Ergodic averaging sequences. *J. Analyse Math.* **95** (2005), 63–103.
[10] J. Bourgain. On the maximal ergodic theorem for certain subsets of the positive integers. *Israel J. Math.* **61** (1988), 39–72.
[11] J. Conze, E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *Probabilités, Publ. Inst. Rech. Math. Rennes*, 1987-1, Univ. Rennes I, Rennes, (1988), 1–31.
[12] L. Corwin and F. P. Greenleaf. Representations of nilpotent Lie groups and their applications Part 1: Basic theory and examples. *Cambridge Studies in Advanced Mathematics* **18** (1990).
[13] N. Frantzikinakis. The structure of strongly stationary systems. *J. Analyse Math.* **93** (2004), 359–388.
[14] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.* **360** (2008), no. 10, 5435–5475.
[15] N. Frantzikinakis. Multiple recurrence and convergence of Hardy sequences of polynomial growth. Preprint 2009. Available at arXiv:0903.0042
[16] N. Frantzikinakis, B. Kra. Polynomial averages converge to the product of integrals. *Isr. J. Math.* **148** (2005), 267–276.
[17] N. Frantzikinakis, B. Kra. Convergence of multiple ergodic averages for some commuting transformations. *Ergodic Theory Dynam. Systems* **25** (2005), no. 3, 799–809.
[18] N. Frantzikinakis, B. Kra. Ergodic averages for independent polynomials and applications. *J. London Math. Soc.* (2) **74** (2006), no. 1, 131–142.
[19] N. Frantzikinakis, E. Lesigne, M. Wierdl. Powers of sequences and recurrence. *Proc. Lond. Math. Soc.* (3) **98** (2009), no. 2, 504–530.
[20] N. Frantzikinakis, M. Wierdl. A Hardy field extension of Szemerédi's theorem. *Adv. Math.* **222** (2009), 1–43.
[21] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* **71** (1977), 204–256.
[22] B. Green, T. Tao. Linear equations in primes. To appear in *Annals of Math.* Available at arXiv:0606088
[23] B. Green, T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. Preprint 2007. Available at arXiv:0709.3562v4
[24] B. Green, T. Tao. The Möbius function is strongly orthogonal to nilsequences. Preprint 2008. Available at arXiv:0807.1736
[25] I. Håland-Knutson. Phd Thesis, Ohio State University, Columbus, (1992). Available at http://home.hia.no/∼ingerjh/forskning/research.htm
[26] G. Hardy. Properties of logarithmico-exponential functions. *Proc. London Math. Soc.* (2) **10** (1912), 54–90.
[27] G. Hardy. Orders of Infinity. *Cambridge Tracts in Math. and Math. Phys.*, 12 (2nd edition), Cambridge, (1924).
[28] B. Host, B. Kra. Nonconventional ergodic averages and nilmanifolds. *Annals of Math.* **161** (2005), 397–488.
[29] B. Host, B. Kra. Convergence of polynomial ergodic averages. *Isr. J. Math.* **149** (2005), 1–19.
[30] B. Host, B. Kra. Uniformity seminorms on ℓ^∞ and applications. *J. Analyse Math.* **108** (2009), 218–276.
[31] R. Jones, M. Lacey, M. Wierdl. Integer sequences with big gaps and the pointwise ergodic theorem. *Ergodic Theory Dynam. Systems* **19** (1999), no. 5, 1295–1308.
[32] L. Kuipers, H. Niederreiter. Uniform distribution of sequences. Pure and Applied Mathematics. *Wiley-Interscience*, New York-London-Sydney, (1974).
[33] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005) no. 1, 201–213.

[34] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Isr. J. Math.* **146** (2005), 303–316.
[35] E. Lesigne. Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. *Ergodic Theory Dynam. Systems* **11** (1991), no. 2, 379–391.
[36] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.* **91** (1969), 757–771.
[37] M. Ratner. Raghunatan's topological conjecture and distribution of unipotent flows. *Duke Math. J.* **61** (1991), no. 1, 235–280.
[38] J. Rosenblatt, M. Wierdl. Pointwise theorems via harmonic analysis. Ergodic theory and its connections with harmonic analysis (Alexandria, 1993). *London Math. Soc. Lecture Note Ser.* **205**, 3–151.
[39] N. Shah. Limit distributions of polynomial trajectories on homogeneous spaces. *Duke Math. J.* **75** (1994), no. 3, 711–732.
[40] E. Vaughan. The Hardy-Littlewood methos. *Cambridge tracts in mathematics*, **125**, 2nd edition, Cambridge University press, Cambridge, (1997).
[41] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20** (2007), 53–97.
[42] T. Ziegler. Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$. *J. Anal. Math.* **99** (2006), 249–266.

(Nikos Frantzikinakis) Department of Mathematics, University of Memphis, Memphis, TN, 38152, USA

E-mail address: `frantzikinakis@gmail.com`